Transcriptional regulation of Taxol^{TM} biosynthesis in *Taxus cuspidata* procambium cells

Thomas Waibel

A thesis submitted for the degree of

Doctor of Philosophy

Institute of Molecular Plant Sciences

School of Biological Sciences

The University of Edinburgh

December 2010

Contents

Declaration **vi**

Acknowledgements **viii**

Publication **xii**

Abstract **xiv**

List of Abbreviations **xxii**

1 Introduction **1**

1.1 Natural products in medicine 1

1.1.1 Cancer drugs of plant origin 4

1.2 Plant secondary metabolism 5

1.2.1 Plant Terpenoids 9

1.3 Jasmonate-responsive gene expression 12

1.3.1 The plant hormone jasmonic acid 12

1.3.2 Jasmonate signal transduction 14

1.3.3 Jasmonate-responsive promoter elements 16

1.4 TF families in secondary metabolism 19

1.4.1 The AP2/ERF family TFs 19

1.4.2 The bHLH and bZIP family TFs 21

1.4.3 Pathogen-responsive WRKY family TFs 22

1.4.4 The MYB TF family 23

1.5 The biology and chemistry of Taxol™ 24

1.5.1 The history of Taxol™ 25

1.5.2 Taxol™ biosynthetic pathway 26

1.5.3 Taxol™ bioactivity 28

1.5.4 Supply and synthesis of Taxol™ 29

1.6 *Taxus cuspidata* - the model plant 30

1.6.1 *Taxus* as plant model 33

1.6.2 Botany of *Taxus* 34

1.7 Project aims 39

2 Materials and Methods **41**

2.1 Plant material 41

2.2 Plant transformation methods 43

2.3 Enzymatic reporter assays 46

2.4 DNA manipulation and analysis 48

2.5 PCR-based methods 55

2.6 Northern blot analysis 61

2.7 Genomic library construction 62

2.8 Web resources 66

3 Cloning and analysis of Taxol™ biosynthetic gene promoters 67

3.1 Introduction 67

3.2 Genomic library construction attempts 69

3.3 Genome walking 76

3.4 In-silico promoter analysis 79

3.5 Functional characterisation of the *TASY* and *DBTNBT* promoters 81

3.6 Discussion 88

4 Transcriptiome sequencing 97

4.1 Sequencing Methods 98

4.1.1 The Roche/ 454 technology 99

4.1.2 The Illumina/ Solexa technology 101

4.2 Experimental design 102

4.3 Transcriptome characterisation 104

4.4 The procambium csc specific transcripts 110

4.5 The JA-responsive transcriptome and *T. cuspidata* TFs 116

4.6 Discussion 121

5 Gene expression profiling of Taxol™ structural genes in CMCs 129

5.1 Introduction 129

5.2 *BAPT* is up-regulated in CMCs 130

5.3 DGE data of Taxol^{TM} structural genes 133

5.4 Discussion 135

6 General discussion **139**

Appendix I **204**

Appendix II **206**

Appendix III **208**

Appendix IV **209**

Appendix V **210**

Appendix VI **220**

Appendix VII **227**

Appendix VIII **228**

7 Publication **230**

Declaration

I hereby declare that the work presented here is my own and has not been submitted in any form for any degree at this or any other university. Any contribution made by other parties is clearly acknowledged.

Thomas Waibel

Acknowledgements

Firstly, I would like to thank BBSRC for funding my PhD.

I would like to thank my first and second supervisors Prof. Gary J. Loake and Dr. Gwyneth Ingram. My thank goes also to all past and present lab members.

For providing the plant material utilised in this thesis I would like to thank the Biotech company Unhwa.

For advice and guidance regarding the particle bombardment and *Agrobacterium*-mediated infiltration of *N. benthamiana* I would like to thank the members of the Oparka lab.

For advice and guidance regarding the protoplast work I would like to thank Dr. Ombretta Foresti and Dr. Jürgen Denecke at the University of Leeds.

Special thanks goes to the Gene-pool genomics facility team and to Florian Halbritter for his invaluable support regarding the statistical analysis of the *T. cuspidata* transcriptome. I also would like to thank my collegues Yan Zejun, Eunjung Kwon and Rabia Amir who helped with aspects of the transcriptome annotation. And, last but not least, I would like to thank Dr. Sophie Haupt and Dr. Sarah Martin.

For my fianceé Imogen who I love,

und für meine Eltern Alfred und Marianne, die mich immer unterstützt haben.

Publication arising from this work:

Lee, E.-K., Jin, Y.-W., Park, J. H., Yoo, Y. M., Hong, S. M., Amir, R., Yan, Z., Kwon, E., Elfick, A., Elfick, Tomlinson, S., Halbritter, F., Waibel, T., Yun, B.-W., and Loake, G. J., (2010) Cultured cambial meristematic cells as a source of plant natural products. Nat Biotechnol 28, 1213-1217.

Abstract

This thesis presents an investigation into the transcriptional regulation of Taxol™ biosynthsis in *Taxus cuspidata* cell suspension cultures. The potent anticancer drug Taxol™ has been shown to be successful in the treatment of breast, lung and ovarian cancer and the acquired immunodeficiency syndrome (AIDS) related Kaposi's sarcoma. Produced by all species of yew, Taxol™ belongs to the class of taxane diterpenoids and is of huge pharmaceutical importance.

The plant material utilised in this thesis is a cell suspension culture initiated from isolated procambium cells of *T. cuspidata*. The latter is a meristematic tissue giving rise to the conductive tissue of plants. This un-differentiated cell suspension culture exhibits an increased and stable production of Taxol™ in response to the plant hormone elicitor methyljasmonate, limited cell aggregation and fast growth when compared to a cell suspension culture initiated from differentiated cells (somatic) of *T. cuspidata*.

In order to assess the stem cell characteristics of the employed procambium cell suspension culture, the transcriptome of *T. cuspidata* was sequenced utilising Roche/ 454 and Illumina/ Solexa *Nla*III tag sequencing technolo-

gies. Statistical analysis uncovered differential expression profiles of 563 genes present within the procambium cell derived transcriptome by comparison with the somatic cell derived transcriptome. Gene ontology analysis of the latter identified that genes associated with response to stress and defence response were upregulated in the differentially expressed portion within the procambium cell suspension culture. This is consistent with the characteristics of animal stem cells which exhibit robust defence strategies to environmental stress. Furthermore *PHLOEM INTERCALATED WITH XYLEM* (*PXY*) and *TRACHEARY ELEMENT DIFFERENTIATION 2* (*TED2*), which are essential for ordered procambium cell division and differentiation into trachaery elements respectively in *A. thaliana* and *Z. elegans*, are up-regulated in the *T. cuspidata* procambium cell suspension culture.

Further *T. cuspidata* homologues of the jasmonate signalling components JASMONATE ZINC FINGER LIKE ZIM DOMAIN 2 (JAZ2) and JAZ3 were identified among up-regulated transcripts in response to jasmonate treatment in both the procambium and the somatic cell line. Blast analysis identified 211 transcription factors within the APETELA 2 (AP2), BASIC-HELIX-LOOP-HELIX (bHLH), WRKY, MYB and BASIC-LEUCIN-ZIPPER (bZIP) families. Further characterisation established 21 transcription factors which are significantly up-regulated in response to jasmonate treatment and show a higher expression level in procambium cells. These provide promising targets for further functional characterisation to elucidate their involvement within Taxol^{TM} biosynthesis.

In order to investigate transcriptional regulation of the Taxol^{TM} struc-

tural genes, a 513 bp fragment corresponding to the *TAXADIENE SYNTHASE* (*TASY*) promoter was cloned by genome walking. In-silico analysis of the *TASY* and *3'-N-DEBENZOYLTAXOL N-BENZOYLTRANSFERASE* (*DBTNBT*) promoter resulted in the identification of methyljasmonate and pathogen-responsive elements which may significantly contribute to jasmonate mediated accumulation of Taxol™. Analysis of a chimeric promoter construct driving the reporter gene *β-GLUCURONIDASE* (*GUS*) in *N. benthamiana* confirmed jasmonate-responsiveness of the *TASY* promoter.

Finally, comparison of the expression level of genes coding for potentially rate-limiting enzymes within the Taxol™ pathway established a significantly increased expression of *BACCATIN II PHENYLPROPANOYLTRANSFERASE* (*BAPT*) in response to jasmonate treatment within the procambium cell suspension culture. Furthermore transcripts of *TASY*, *PHENYLALANINE AMINOMUTASE* (*PAM*) and *DBTNBT* show an overall higher expression and prolonged transcript accumulation in procambium compared to somatic cells.

In this thesis jasmonate-signalling components, jasmonate-responsive transcription factors and differential gene expression profiles of Taxol™ structural genes were identified which, may contribute to an increased Taxol™ production in the utilised procambium cell suspension culture. Furthermore the *T. cuspidata* procambium cell suspension culture was found to have an increased level of stress- and defence-response reflected by differential gene expression profiles and content of phenolic compounds and Taxol™.

List of Figures

1.1 Groups and function of secondary metabolites 8

1.2 Terpenoid formation 11

1.3 Jasmonate signal transduction 15

1.4 Taxol^{TM} biosynthetic pathway 27

1.5 Cross section 31

1.6 *Taxus cuspidata* 36

1.7 *Taxus* habitat 37

1.8 Model organism divergence 38

2.1 Reporter constructs 50

2.2 pJIT166 and pGreen0049 50

2.3 Genome walking adapter 57

3.1 *Hind*III test restriction 70

3.2 Size selection 71

3.3 HMW DNA and pIndigoBAC-5 for BAC library construction . . 73

3.4 BAC insert size determination 75

3.5 Genome walking I 77

3.6 Genome walking II 78

3.7 Particle bombardment 82

3.8 GUS activity assay on biolastic transformed cells 83

3.9 α-amylase activity assay 86

3.10 *TASY* and *DBTNBT* promoter and the Gateway® vector pMDC162 88

3.11 GUS activity assay 89

4.1 Roche/ 454 methodology 100

4.2 Experimental design 103

4.3 RT-PCR MeJA elicitation 104

4.4 Plant species of contig annotation 105

4.5 Plant species of contig annotation 107

4.6 Solexa tag count 109

4.7 Heatmap d.e. contigs in CMCs 111

4.8 GO term enrichment 112

4.9 Alignment of contig 01805 113

4.10 DGE contigs 01805 and 13935 114

4.11 *TED2* degenerate PCR and northern analysis 115

4.12 Heatmap d.e. contigs in after MeJA treatment 118

4.13 Identified TFs 119

5.1 Jasmonate induction of Taxol™ structural genes 130

5.2 *TASY* and *PAM* ROI analysis 132

5.3 *BAPT* and *DBTNBT* ROI analysis 133

5.4 DGE expression profiles of Taxol™ structural genes 134

6.1 Taxol™ pathway regulation in CMCs 141

6.2 TE differentiation stages 145

6.3 Taxol™ pathway branches 150

6.4 TED2 alignment 204

6.5 AP2/ERF TFs expression profiles I 220

6.6 AP2/ERF TFs expression profiles II 221

6.7 bHLH TFs expression profiles 222

6.8 WRKY TFs expression profiles 223

6.9 MYB TFs expression profiles I 224

6.10 MYB TFs expression profiles II 225

6.11 bZIP TFs expression profiles 226

List of Tables

1.1 Cancer drugs of plant origin 3

1.2 Uses of secondary metabolites 6

1.3 Plant terpenoid classification 10

1.4 Essential G-boxes 18

2.1 *T. cuspidata* medium composition 42

3.1 *TASY* promoter *cis* elements 80

3.2 *DBTNBT* promoter *cis* elements 81

4.1 Assembly and annotation statistics 106

4.2 Sequence identities 108

6.1 Mixed base definitions 205

List of Abbreviations

4-MU	4-methylumbelliferone
10-DAB	10-deacetylbaccatin III
AIF	ATBS1-interacting factor
AP2	APETELA 2
as-1	activation sequence-1
AtBS1	ACTIVATION TAGGED BRASSINOSTEROID-INSENSITIVE1 SUPPRESSOR
BAC	bacterial artificial chromosome
BAPT	BACCATIN II PHENYLPROPANOYLTRANSFERASE
BC	Before Christ
Bcl2	B-cell lymphoma 2
bHLH	basic-helix-loop-helix
BR	brassinosteroid
bZIP	basic leucine zipper
C1	COLORLESS
CaM V	Cauliflower mosaic virus
CCNSC	Cancer Chemotherapy National Service Center
cDNA	copy DNA
cfu	colony forming unit
CLE	CLAVATA/ESR1-LIKE
CLV	CLAVATA
CMCs	cambial meristematic cells

CoA	coenzyme A
COI1	CORONATINE INSENSITIVE1
csc	cell suspension culture
DBAT	10-DEACETYLBACCATIN III-10-O-ACETYL TRANSFERASE
DBBT	2-DEBENZOYL-7,13-DIACETYLBACCATIN III-2-O-BENZOYL TRANSFERASE
DBTNBT	3'-N-DEBENZOYLTAXOL N-BENZOYLTRANSFERASE
DDCs	dedifferentiated cells
d.e.	differentially expressed
DGE	digital gene expression
DMAPP	dimethylallyl pyrophosphate
DMPP	dimethylallyl pyrophosphate
DNA	deoxyribonucleic acid
DXP	1-deoxy-D-xylulose 5-phosphate
EREBP	ethylene-responsive element binding protein
ERF	ETHYLENE RESPONSE FACTOR
EST	expressed sequence tag
FCW	fresh culture weight
FPP	farnesyl diphosphate
G2/M	Gap2/ Mitosis
GDP	guanosine diphosphate
GFP	GREEN FLUORESCENT PROTEIN
GGPP	geranylgeranyl pyrophosphate
GGPPS	GERANYLGERANYL PYROPHOSPHATE SYNTHASE
GPP	geranyl pyrophosphate
GTP	guanosine triphosphate
GO	gene ontology
GUS	β-glucuronidase
HIV	Human immunodeficiency virus
HMW	high molecular weight
H-NMR	hydrogen nuclear magnetic resonance

HPLC	high pressure liquid chromatography
Ile	isoleucine
IPP	isopentenyl pyrophosphate
IUCN	International Union for Conservation of Nature
JA	jasmonic acid
JAK-STAT	JANUS KINASE/ SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION
JAR1	*JASMONATE RESISTANT1*
JAZ	JASMONATE ZINC FINGER LIKE ZIM DOMAIN
JERE	jasmonate and elicitor response element
LAP	*LEUCINE AMINOPEPTIDASE PUTRESCINE*
MeJA	methyl jasmonate
MEP	2-C-methyl-D-erythritol 4-phosphate
MUG	4-methylumbelliferyl β-D-glucuronide
MVA	mevalonic acid
MYA	million years ago
MYB	from v-myb: avian myeloblastosis virus
MYC	from c-myc similar to v-myc: myelocytomatosis viral oncogene
NCI	National Cancer Institute
NINJA	NOVEL INTERACTOR OF JAZ
*Nla*III	*Neisseria lactamica* III
OD	optical density
OPP	diphosphate
ORCA	OCTADECANOID-DERIVATE RESPONSIVE *CATHARANTHUS* AP2-DOMAIN PROTEIN
ORF	open reading frame
p53	protein 53
PAM	PHENYLALANINE AMINOMUTASE
PCF	plant cell fermentation
PCR	Polymerase chain reaction
PDF	PLANT DEFENSIN GENE
PFG[E]	pulse field gel [electrophoresis]
PR	*PATHOGEN RELATED*
PTI	PTO-INTERACTING PROTEIN

PTO	*Psydomonas syringone* pv tomato
PP_i	pyrophosphate
PXY	*PHLOEM INTERCALATED WITH XYLEM*
R2R3	repeat 2 repeat 3
R	RED
RNA	ribonucleic acid
ROI	region of interest
RT-PCR	reverse Transcriptase PCR
SCF	SKIP-CULLIN-F-box
SCLC	small cell lung cancer
scv	setteld cell volume
STR	*STRICTOSIDINE SYNTHASE*
T5αH	taxadiene 5alpha-hydroxylase
TαH	TAXANE 13-ALPHA HYDROXYLASE
T10βH	5-ALPHA-TAXADIENOL-10-BETA-HYDROXYLASE
TASY	TAXADIENE SYNTHASE
TDAT	TAXADIENOL ACETYL TRANSFERASE
T-DNA	transfer DNA
TE	Tracheary elements
TF	transcription factor
TED2	*TRACHEARY ELEMENT DIFFERENTIATION2*
TIA	terpenoid indole alkaloid
TPL	TOPLESS
tRNA	transfer RNA
U	Unit
USDA	United States Department of Agriculture
UTR	untranslated region
UV	ultra violet
V	volt
WHO	World Health Organisation
WRKY	from WRKYGQK consensus
ZIM	ZINC-FINGER PROTEIN EXPRESSED IN INFLORESCENCE MERISTEM

Glossary

B16 line:	Mouse skin melanoma cell line.
cis element:	From the Latin cis, meaning "on the same side as". A *cis*-element refers to a short regulatory sequence upstream of the transcription start site of a gene which serves as a binding site for trans-acting "acting from a different molecule" TFs.
homogenous cells:	The characteristics of all cells is uniform.
heterogenous cells:	The characteristics of all cells is not uniform.
L1210 and P388 lines:	Mouse lymphocytic leukemia cell lines, extensively used in cytotoxicity assays of chemotherapeutic agents.

Chapter 1

Introduction

1.1 Natural products in medicine

Plants have always been of major importance to mankind, serving as food, shelter, clothing, fuel, tools and many more. As an integral part of every culture throughout history, plants have also played a vital role in traditional medicine in treating and preventing human diseases. The earliest written account of drugs of plant origin date from about 2600 BC and were found in Mesopotamia. Oils of *Cedrus* species and cypress (*Cupressus sempervirens*) are recorded, both of which are still in use today [Gurib-Fakim, 2006]. Early medicines, besides oil, were administered in the form of crude extracts, teas, poultices or powders [Balunas and Kinghorn, 2005].

In more recent history the first drugs in sophisticated western medicine were also predominantly of plant origin; namely aspirin, digitoxin, morphine, quinine and philocarpine [Butler, 2005]. The importance of phytopharmaca

continues in current times and is highlighted by 28% of new drugs approved between 1981-2002 being plant natural products with a further 24% synthesised around a pharmacore derived from a plant compound [Harvey, 2007]. Furthermore the WHO classified 252 drugs as basic and essential from which 11% are produced by plants [Rates, 2001]. The great success of drugs from plant origin can be ascribed to a number of characteristics: (1) Plant natural products have a high structural diversity with common structural features, which have been shown to be highly relevant in drug discovery such as chiral centres, aromatic rings, complex ring systems, degree of molecule saturation and number and ratio of heteroatoms [Balunas and Kinghorn, 2005], (2) the high bioactivity of plant-derived compounds can be attributed to structural similarity of protein targets across many species, (3) conventional medicine can be inefficient (e.g. side effects) and (4) a growing ecological awareness suggests that plant-derived natural medicine is harmless [Harvey, 2007].

Furthermore plant products can be utilised as lead-compounds important in the development and synthesis of new drugs [Rates, 2001]. There is a huge potential for the discovery of new relevant plant natural products with therapeutic activity, which are mainly synthesised in plant secondary metabolism [reviewed in 1.2] since only 5% - 15% of higher plants have been chemically analyzed [Wink, 1985]. This constitutes a big opportunity for the identification of novel chemical classes with a potentially undiscovered mechanism of action or cellular targets relevant to treatment and drug design [Lindholm, 2005].

Cancer drugs of plant origin.

Name	Origin	Therapeutic use	Publication	Mechanism of action
Taxol^{TM}	*Taxus brevifolia*	Ovarian, breast and non-small lung cancer	[Wani et al., 1971]	Microtubule stabilisation
Docetaxel	*Taxus baccata*	Ovarian, breast and bronchial carcinomas	[Mangatal et al., 1989]	Microtubule stabilisation
Camptothecin	*Camptotheca acuminata*	Gastrodintestinal tumors	[Wall et al., 1966]	Enhances binding of topoisomerase I to DNA, thus promoting DNA strand breaks
Irinotecan	Synthesised from camptothecin	Colorectal carcenoma	[Kunimoto et al., 1987]	DNA topoisomerase I inhibitor
Topotecan	Analogue of camptothecin	Ovarian carcenoma	[Kingsbury et al., 1991]	DNA topoisomerase I inhibitor
Etopside	*Podophyllum sp.*	SCLC, leukemia, non-Hodgkin lymphoma, Hodgkińs disease and testicular cancer	[Keller-Juslen et al., 1971]	DNA topoisomerase II inhibitor
Teniposide	related to Etoposide	Used in combination therapies	[Staehelin, 1970]	DNA topoisomerase II inhibitor
Vinblastine	*Vinca rosea*	Hodgkińs disease and choriocarcinoma	[Neuss et al., 1962] [Neuss et al., 1964] [Neuss et al., 1959]	Antimitotic, inhibition of cell division
Vincristine	Vinca rosea	Leukemia, SCLC and malign lymphoma	[Farnsworth, 1988]	Antimitotic
Vinorelbine	Synthesised from vinblastine	SCLC and breast cancer	[Mangeney et al., 1979]	Antimitotic, inhibiton of tubulin polymerisation

Table 1.1: The table lists compounds of plant origin or their analogues which are used in cancer treatment. A common target is tubulin resulting in failure of proper microtubule dynamics thus hampering cell division [reviewed in 1.4.3] and topoisomerase I and II. Cancer drugs targeting the latter promoting topoisomerase generated DNA breaks [Pommier, 1993, Baldwin and Osheroff, 2005]. The drugs can be grouped into taxenes: Taxol^{TM} and docetaxel; camptothecins: camptothecin, irinotecan and topotecan; epipodophyllotoxons: etopside and teniposide and Vinca alkaloids: vinblastine, vincristine and vinorelbine.

1.1.1 Cancer drugs of plant origin

The demand for anticancer drugs is bigger than ever with cancer being the second leading cause of death with roughly 600,000 in the United States (recorded in 2007) and over 1.4 million new cases of cancer diagnosed annually [Jemal et al., 2007]. The most common forms are lung, breast, colorectal and stomach cancer, which showed an increase of 22% since 1990 [Parkin et al., 2001].

Drugs of plant origin or mimics of natural products accounted for 48% of the drugs available in the treatment of cancer between 1940 and 2002 [Newman et al., 2003]. The main anticancer drugs of plant origin [Table 1.1] are camptothecins from *Campotheca acuminata* [van der Heijden et al., 2004], Vinca alkaloids (vinblastine and vincristine) from *Catharanthus roseus* (Madagascar periwinkle) [Okouneva et al., 2003], epipodophyllotoxins from *Podophyllum peltatum* and taxenes including paclitaxel or Taxol^{TM} originally isolated in *Taxus brevifolia* but sythesised in all species of *Taxus* [Balunas and Kinghorn, 2005].

One third of anticancer drug sales in 2002 can be accredited to taxenes and camptothecins with a combined total of over 2.75 billion dollars [Oberlies and Kroll, 2004]. Although of huge pharmaceutical and economical importance, the vast majority of anticancer drugs of plant origin cannot be synthesised economically or rely on an unsustainable semi-synthetic process with precursors of plant origin [Rates, 2001]. This highlights the need for in-depth molecular knowledge of plants such as *Taxus* , thus providing promising targets for metabolic engineering.

1.2 Plant secondary metabolism

Integral to the life cycle of plants is their ability to protect themselves against abiotic- and biotic-stress. The coexistence of plants and insects for 350 million years have triggered the development of a diverse array of plant defensive mechanisms and contributed to increased diversity of plant species, defensive structures and herbivors mode of attack [Gatehouse, 2002]. The group of compounds responsible for a plethora of chemical protective measures against herbivores and abiotic-stress are collectively termed secondary metabolites [Pichersky and Gang, 2000].

In the past few decades the interest in secondary metabolites has increased as more of this diverse group of plant products find their way into our modern life [Wink, 1985]. We exploit them as pharmaceuticals (e.g. the diterpenoid Taxol^{TM}), fragrances, flavours and many more [Table 1.2]. Thus knowledge about their biosynthesis is of huge biotechnological and economical importance [Wink, 1985].

By definition secondary metabolites are not essential for growth and development but indespensable for the survival of a plant [Hartman, 1996]. Secondary also in the sense that their biosynthesis is restricted to a specific plant group [Pichersky and Gang, 2000]. Primary metabolism is conserved and uniform among the plant kingdom, by contrast secondary metabolites are diverse and adaptive, contributing to the interaction of the plant with its environment [Fig.: 1.1] [Hartman, 1996]. The diversity of secondary metabolites is demonstrated by the existence of subtly different defence strategies among different genotypes of the same plant species [Jander et al., 2001]. The level of speci-

Examples of secondary metabolites uses in modern life.

Use	Compound	Plant species
Pharmaceutical	Taxol™	*Taxus brevifolia*
	Atropine	*Atropa belladonna*
	Scopolamine	*Datura stramonium*
	Quinine	*Cinchona pubesence*
	Cardenolides	*Digitalis latana*
	Codeine	*Papaver somniferum*
Fragrances	Rose oil	*Rosa sp.*
	Lavender oil	*Lavendula angustifolia*
Flavours	Vanillin	*Vanilla planifolia*
	Capsaicin	*Capsicum frutenscens*
Colours	Indigo	*Indigofera tinctoria*
	Shikonine	*Lithospermum erythrorhyzon*
Poisons	Coniine	*Conium maculatum*
	Strychnine	*Strychnos nux vomicaend*
Stimulants	Caffeine	*Coffea arabica*
	Theophylline	*Thea sinensis*
	Nicotine	*Nicotiana sylvestris*
Hallucinogen	Cocaine	*Erythroxylon coca*
	Cannabinol	*Cannabis sativa*
Insecticide	Nicotine	*Nicotiana sylvestris*
	Pyrethrin	*Pyrethrum cinerariifolium*

Table 1.2: The table gives examples of plant secondary metabolites exploited by man for various purposes with the compound name and the plant species of origin [Wink, 1985].

ficity enables the use of such compounds as taxonomic markers [Bennett et al., 1994]. The chemical composition present in *Taxus cuspidata*, the plant model used for this thesis, was found to be very different from other species within the same genus but most resembled that of *Taxus canadensis* [Wang et al., 2010].

Foundation of the plant specific array of secondary compounds are key-intermediates from which a few biosynthetic routes lead to numerous products as seen among monoterpene biosynthesis [Hartman, 1996]. A further contributing factor is found among the biosynthetic enzymes involved. Some have been shown to synthesise multiple products from different, or the same, substrates [Allina et al., 1998, Maury et al., 1999, Phillips et al., 1999].

From their original proposed function as protectants against biotic- and abiotic-stress, secondary metabolites have evolved into other biological activities such as attraction of pollinators or allelopathy [Fig.: 1.1] [Wink, 1985]. It is noteworthy that some compounds combine the function of repellent and attractant. For example volatile terpenes function as antimicrobials but also attract pollinators [Wink, 1985].

Among the secondary metabolites accumulated for plant defence are two classes which can be characterised as static and active. The first group includes compounds constitutively produced and contributing to plant defence such as physical barriers, resin, deterrents of feeding or egg deposition or toxins [Bennett et al., 1994, Gatehouse, 2002]. Should a herbivore attack, the accumulated and stored static compounds provide the necessary level of protection [Gatehouse, 2002]. In contrast active metabolites are only accumulated

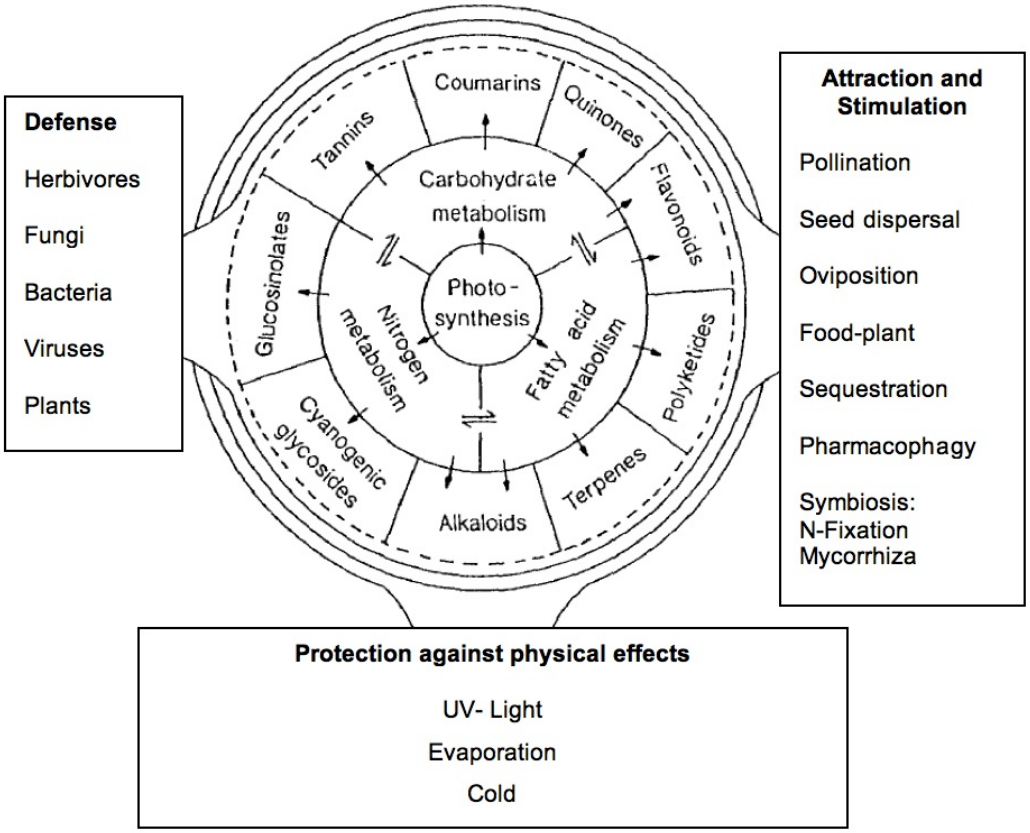

Figure 1.1: Secondary metabolite biosynthesis from common precursors of primary metabolism and their functions within the interaction of a plants with its environment [Hartman, 1996]

in response to biotic- or abiotic-stress as their on-demand reversible biosynthesis constitutes resource limitation. For example this becomes apparent in *Nicotiana sylvestris* where the nitrogen employed for the synthesis of the alkaloid nicotine can't be recycled into the plant metabolism and can therefore be considered wasted [Baldwin, 1998]. A drawback in the induced accumulation of chemical defence compounds is the presence of a lag time from the external stimulus until the release of secondary metabolites, thereby allowing

biotic- or abiotic-stresses to inflict potentially irreparable damage [Baldwin, 1998, Agrawal et al., 2002]. Taxol^{TM} also belongs to the class of active secondary metabolites, showing accumulation in response to biotic elicitors [Nims et al., 2006].

1.2.1 Plant Terpenoids

Taxol^{TM} belongs to the class of taxane diterpenoids, more specifically to the taxoids with an oxetane ring and a β-phenylalanine C-13 side chain. Increasing interest in Taxol^{TM} has triggered the isolation and characterisation of over 360 taxane diterpenoids within 16 subgroups from various plant parts of *Taxus* species, most of which contain the pentamethyl [9.3.1.0]^{3,8} tricyclopentadecane (taxoid) skeleton [Baloglu and Kingston, 1999, Jennewein et al., 2003].

With over 30,000 identified compounds, plant terpenoids form the largest class of natural products [Theis and Lerdau, 2003, Connolly and Hill, 1991]. Terpenes are found in primary metabolism as hormones (e.g. gibberellins, abscisic acid), photosynthetic pigments (e.g. carotenoids), electron carriers (e.g. ubiquinone), mediators of polysaccharide assembly (e.g. polyprenyl phosphates) and structural membrane components (e.g. phytosterols) [Harborne, 1991]. Others function as secondary products including herbivore deterrents, fungal toxicity and pollinator attraction [Theis and Lerdau, 2003].

Plant terpenoids are synthesised from the repetitive joining of IPP (isopentenyl pyrophosphate) and can hence be classified according to the number of IPP molecules utilised, increasing sequentially by five carbon units [Table 1.3] [McGarvey and Croteau, 1995, Bouvier et al., 2005]. IPP itself can be syn-

Plant terpenoid classification

Terpenoid class	number of IPP units	number of carbon units
Hemiterpenoids	1	5
Monoterpenoids	2	10
Sesquiterpenoids	3	15
Diterpenoids	4	20
Sesterpenoids	5	25
Triterpenoids	6	30
Tetraterpenoids	8	40
Polyterpenoids	≤ 8	NA

Table 1.3: The table shows the classification of plant terpenoids according to the number of IPP and carbon units present.

thesised in the cytosol and the plastids via the MVA or the MEP (sometimes refered to as DXP) pathway respectively [Bouvier et al., 2005]. Although terpenoids constitute a large group of compounds, they are all synthesised via a common pathway [Fig.: 1.2] thus complex regulation must be in place to ensure their correct spatial and temporal distribution at an appropriate level [McGarvey and Croteau, 1995].

Biosynthesis of terpenoids often occurs in a tissue-specific manner. Monoterpenes designated for plant defence within species of *Mentha* (mint) are synthesised in glandular trichomes on the leaf surface [Gershenzon et al., 1989]. Linalool utilised by *Clarkia breweri* (Brewer's clarkia) to attract pollinators is synthesised in the flower petals [Pichersky et al., 1994]. In conifers the synthesis of wound induced mono- and diterpenoids is restricted to the area of stimulus [Blanchette, 1992].

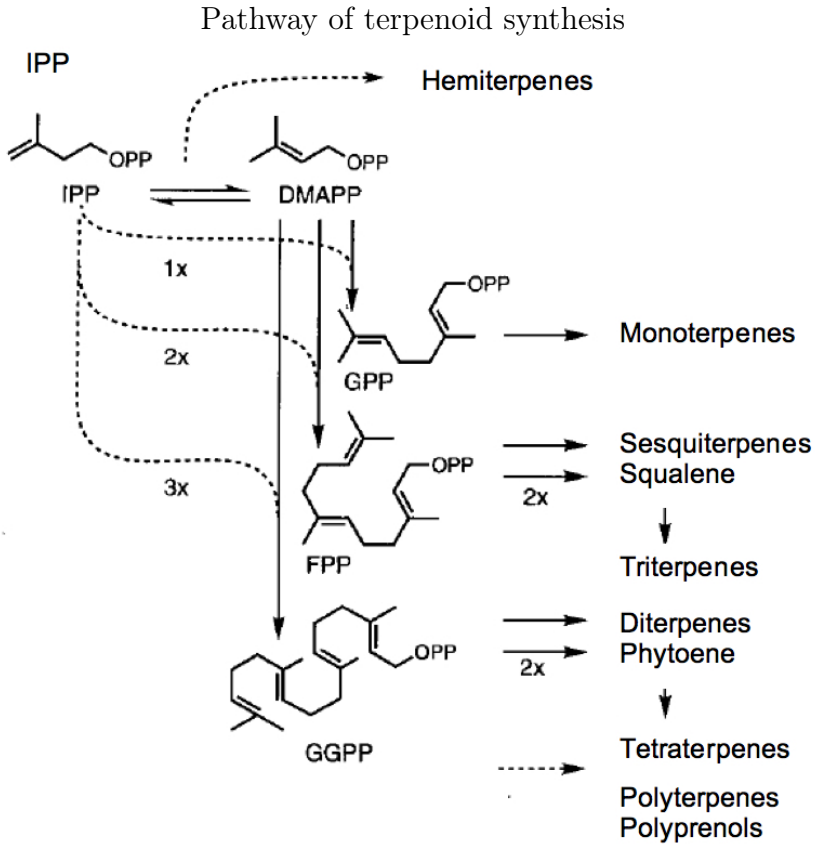

Figure 1.2: Schematic outlining the formation of different terpenoids from common precursors. Source: [McGarvey and Croteau, 1995]

Diterpenoids are synthesised in the plastids via the MEP pathway as is the utilised IPP, the precursor molecule for Taxol^{TM}, starting from pyruvate and glyceraldehyde-3-phosphate [Kleining, 1989]. TAXADIENE SYNTHASE, the enzyme resposible for the committing step in the Taxol^{TM} biosynthesis pathway [reviewed in 1.4.2.] was found to harbour an *N*-terminal plastidial targeting sequence [Williams et al., 2000].

1.3 Jasmonate-responsive gene expression

Jamonates are key signalling molecules involved in defence response to biotic- or abiotic-stress leading to the accumulation of protective secondary metabolites including Taxol^{TM} [reviewed in 1.2] [McConn et al., 1997, Wu and Ge, 2004]. Biosynthesis of secondary metabolites in response to jasmonates has been shown in a variety of different plant species [Memelink et al., 2001].

The ability of cells to rapidly respond to environmental stimuli depends on the regulation of gene expression. Gene transcription in eukaryotic species is carried out by the RNA POLYMERASE II machinery and the basal transcription factors [reviewed by [Selth et al., 2010]]. However, it is becoming increasingly clear that expression of inducible genes and thus their transcriptional control is governed by more complex systems [reviewed by [Weake and Workman, 2010]]. For example overlap of signal transduction and transcription regulation has been shown in *S. cerevisiae* [Weake and Workman, 2010]. In plants, promoter elements conferring jasmonate-responsiveness such as the GCC-motif and G-box [reviewed in 1.3.3] are dependent on the jasmonate signalling component COI1 [Benedetti et al., 1995, Lorenzo et al., 2003].

1.3.1 The plant hormone jasmonic acid

Jasmonic acid and its methyl ester belong to a class of plant hormones collectively called jasmonates which are synthesised from the oxygenated fatty acid derived precursor linolic acid (18:3) via the octadecanoid pathway [Sembdner and Parthier, 1993, Somerville and Browse, 1991]. α-Linolenic acid is released

through hyrdolisis from choloroplast membranes. Hydrolases associated with jasmonate biosynthesis are the wound-induced phospholipase A_2 and a DEFENCE IN ANTHER DEHISCENCE (DAD)-like phospholipase A_1 [Ishiguro et al., 2001, Narvaez-Vasqueza et al., 1999]. The biosynthesis of jasmonates is reviewed by Schaller et al. [2005].

Biological functions originally associated with the free acid and the methyl ester are growth inhibition and promotion of senesence respectively [Ueda and Kato, 1980, Dathe et al., 1981]. Further developmental functions of jasmonates are inhibition of seed germination, stimulation of tendril coiling, various aspects of flower development, tuberization in potato and others [Wasternack, 2007, Dathe et al., 1981, Corbineau et al., 1988, Ueda and Kato, 1980, Falkenstein et al., 1991, Weiler et al., 1998]. Jasmonates are also involved in response to biotic- and abiotic-stress. JA concentration increases locally in response to pathogen infection or wounding in concert with expression of *PR* genes, which is also shown for exogenously applied JA [Lorenzo and Solano, 2005, Wasternack, 2007, Stintzi et al., 1993]. Expression of proteinase inhibitors and antimicrobial phytoalexins in response to jasmonates highlights their crucial role in plant defence [Farmer and Ryan, 1990, Gundlach et al., 1992]. Defence-gene-induced enhanced resistance against herbivor and pathogens is also mediated by jasmonates [Reymond and Farmer, 1998]. Furthermore mutations in the *A. thaliana* loci *coi1* and *jar1*, both coding for JA signalling components, show compromised defence-gene expression in response to pathogen challenge [Xie et al., 1998, Staswick and Tiryaki, 2004].

1.3.2 Jasmonate signal transduction

COI1, an F-box protein is central to jasmonate signal transduction [Fig.: 1.3] [Xie et al., 1998]. Although identified in *A. thaliana*, homologues exist in *N. attenuata* and *L. esculentum* indicating that the jasmonate signalling is conserved in other species [Paschold et al., 2007, Li et al., 2004]. COI1 forms part of the SCF multiprotein complex which functions as E3 ubiquitin ligase [del Pozo and Estelle, 2000, Devoto et al., 2002]. As such the SCF^{COI1} complex catalyses the ubiquitination of proteins destined for degradation by the 26S proteasome, the F-box of COI1 confers the substrate specificity [Devoto et al., 2002, Xu et al., 2002]. In the presence of the plant hormone, the SCF^{COI1} complex targets the JAZ proteins for degradation [Chini et al., 2007]. In *A. thaliana* 12 members of the JAZ family have been identified which act as JA-repressors [Thines et al., 2007]. It was recently shown that JAZ proteins recruit the Groucho/Tup1-type co-repressor TPL mediated by the adaptor protein NINJA [Pauwels et al., 2010]. Characterisation of the enzyme encoded by the JAR1 locus provided further insights into jasmonate signalling, which was found to conjugate the amino-acid isoleucine (Ile) to JA generating JA-Ile. An interaction between JAZ1 and COI1 in the presence of JA-Ile was subsequently shown which initiates proteolysis mediated degradation of JAZ and activation of the bHLH transcription factor MYC2 [Staswick and Tiryaki, 2004, Lorenzo et al., 2004, Chini et al., 2007]. MYC2 in turn activates expression of early JA-responsive genes including JAZ proteins resulting in a negative feedback loop [Chini et al., 2007].

Jasmonate-responsive induction of secondary pathways in different plant

species occurs in a time specific manner. Transcripts of biosynthetic genes of phenylpropanoid biosynthesis in *Medicago truncatula* peak after 2 hours. In contrast expression of structural genes involved in triterpene biosynthesis in the same species occurs within 12-24 hours [Suzuki et al., 2005]. In tobacco cells the peak of expression of genes involved in nicotine biosynthesis occurs 1-2 hours after jasmonate perception [Goossens et al., 2003]. Hence different

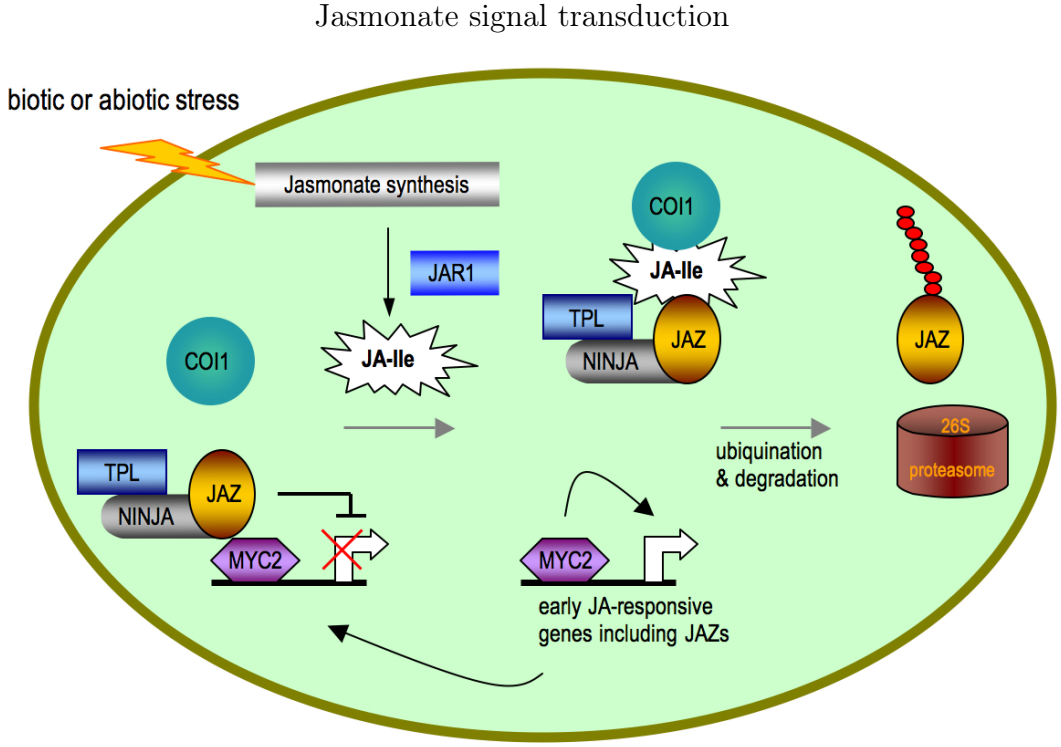

Figure 1.3: Jasmonate signal transduction. Jasmonate biosynthesis is triggered by biotic or abiotic stress. JAR1 generates JA-Ile which increases affinity of the SCF^{COI1} to the JAZ repressor complex. The SCF^{COI1} complex catalyses ubiqitination and degradation of JAZ which forms with TPL and NINJA a repressor complex. Liberation of MYC2 triggers transcriptional activation of early jasmonate-responsive genes [adapted from [Baker et al., 2010]].

modes of downstream regulation after perceiving the jasmonate signal must be in place which result in the observed staggered induction of secondary pathways. However, these may share common components [Pauwels et al., 2009].

1.3.3 Jasmonate-responsive promoter elements

The promoter region of a gene is the controlling element conferring tightly regulated expression in a spatial and temporal dependent manner. Cellular processes are activated through correct gene expression which is also true for response to environmental cues resulting in dynamic signalling cascades and biosynthesis of secondary metabolites [Yamaguchi-Shinozaki and Shinozaki, 2005].

Cis acting elements within gene promoters function as binding sites for *trans* TFs, acting as transcriptional activators or repressors [Latchman, 1997]. These elements are conserved and in combination with several *cis*-elements present in multiple gene promoters allow for specific and co-ordinate gene expression [Singh, 1998].

The GCC promoter element

The GCC motif (GCCGCC), present in several defence-responsive genes including *THIONIN2.1* and *PR* genes was originally identified as an ethylene-responsive element and shown to interact with TFs of the ERF family [Gu et al., 2002]. However it is also present in the promoter of *PLANT DE-*

FENSIN1.2, a common indicator for jasmonate-dependent defence responses [Brown et al., 2003]. The expression of *PR* genes in tobacco is induced by ethylene through the essential GCC promoter element [Ohme-Takagi and Shinshi, 1995]. However, pathogen-responsive jasmonate and ethylene-signalling are both required for expression of *PDF1.2* [Penninckx et al., 1998].

The GCC motif is also part of the a jasmonate-responsive element, JERE, present in the promoter of *STR* which is involved in indole alkaloid biosynthesis in *Catharanthus roseus* [Menke et al., 1999]. JA-responsiveness is abolished by deletion or mutation of the JERE and jasmonate-responsiveness was confirmed by expression of a reporter gene driven by JERE [Menke et al., 1999]. The two TFs ORCA2 and ORCA3 interact with JERE. The latter belong to the AP2/ERF family of TF [reviewed in 1.4] and as such harbour a AP2/ ERF DNA-binding domain which is unique to plants [van der Fits and Memelink, 2000].

The G- and T/G-box promoter element

The G-box promoter element (CACGTG) is the target binding site of MYC2 in jasmonate signal transduction which highlights the importance of this sequence and the related T/G-motif (AACGTG) in jasmonate-responsive transcription [Chini et al., 2007]. G- and T/G-boxes essential for JA-inducibility are present in several promoters [Table: 1.4]. Essential supplementary sequences adjacent to the G-box have been identified in the *LAP* and *ORCA3* gene promoters where a GAGTA repeat and an A/T rich sequence confers correct transcription activation in response to JA and expression level respec-

G-boxes essential for jasmonate-responsive transcription.

Gene name	Plant species	Publication
PROTEINASE INHIBITOR 2	Potato	[Kim et al., 1992]
VEGETATIVE STORAGE PROTEIN B	*Glycine max*	[Mason et al., 1993]
VEGETATIVE STORAGE PROTEIN 1	*A. thaliana*	[Guerineau et al., 2003]
LEUCINE AMINOPEPTIDASE	*L. esculentum*	[Boter et al., 2004]
PUTRESCINE N-METHYLTRANSFERASE 1A	*Podophyllum sp.*	[Xu and Timko, 2004]
ORCA3	*C. roseus*	[Endt et al., 2007]

Table 1.4: The table lists genes from several plant species which harbour in their promoter a G-box motif essential for jasmonate inducibility.

tively [Boter et al., 2004, Endt et al., 2007]. Analysis of *A. thaliana* promoters established a statistically significant overrepresentation of G-box motifs among jasmonate-responsive genes [Mahalingam et al., 2003].

The as-1-type and JASE promoter element

The promoter of *LIPOXYGENASE 1* in barley harbours an as-1-type (TGACG) motif, which is essential for JA-responsiveness [Rouster et al., 1997]. The as-1-type motif is also present and essential in the *NOPALINE SYNTHASE* promoter of the *Agrobacterium thumefaciens* T-DNA [Kim et al., 1993]. Elements conferring jasmonate-responsiveness without a common consensus sequence found in the promoter of *12-OXO-PHYTODIENOIC ACID-10,11-REDUCTASE* are the jasmonate-responsive elements 1 (CGTCAATGAA) and 2 (CATACGTCGTCAA). It is noteworthy that the latter has a ACGT core also present in the G-box motif [He and Gan, 2001].

1.4 TF families in secondary metabolism

Coordinate transcriptional control of pathway genes is key to biosynthesis and accumulation of secondary metabolites [Zhao and Last, 1996, Pelletier et al., 1999]. In the biosynthethis of benzylisoquinoline alkaloids in poppy cell cultures, genes involved in other aspects of biosynthesis including supply of precursors, co-factors and energy molecules (e.g. adenosine triphosphate) required for metabolic enzymes, also show correlated expression to pathway genes [Zulak et al., 2007]. Transcriptional regulation can be attributed to specific TFs which govern the expression of target genes [Endt et al., 2007].

Transcription factors are characterised as modular proteins, harbouring a DNA-binding domain and an autonomous activator or repressor domain. Through interaction with *cis* elements within a target gene promoter in a sequence specific manner, TFs are capable of modulating the rate of transcription [Latchman, 1997]. Transcription factor activity may depend on the interaction with additional co-factors (activators, repressors and in some cases cromatin remodelling factors) or competition with other TFs [Davies and Schwinn, 2003, Singh, 1998].

1.4.1 The AP2/ERF family TFs

The AP2/ERF domain TF family mediate transcriptional activation in response to biotic- and abiotic-stress and jasmonates. The AP2 domain was originally implicated in stress response on three TFs (PTI-4, PTI-5 and PTI-6) which showed interaction with the tomato disease resistance protein PTO

mediating *PR* gene expression [Zhou et al., 1997]. The AP2/ERF domain TFs ORCA2 and ORCA3 are involved in transcriptional regulation of terpenoid indole alkaloid (TIA) biosynthesis in *Catharanthus roseus*. The jasmonate and pathogen-responsive ORCA2 mediates expression of *STR* [Menke et al., 1999]. In contrast ORCA3 induces coordinate expression of *STR*, *TRYPTOPHAN DECARBOXYLASE* and *CYTOCHROME P450 REDUCTASE* in response to jasmonate [van der Fits and Memelink, 2000]. Overexpession of *ORCA3* resulted in increased TIA accumulation [van der Fits and Memelink, 2000].

The AP2/ERF domain TF family can be further classified into an AP2/ERF group, which have only one AP2 domain and function in biotic and abiotic stress response, and an AP2 group which exhibit two AP2 domains and are implicated in cambial tissue development [Riechmann and Meyerowitz, 1998], [van der Graaff et al., 2000]. Although thought to be plant specific TFs of the AP2/ERF family have been identified in certain species of ciliates, cyanobacteria, viruses and *Bacteriophages* [Shigyo et al., 2006]. The AP2/ERF domain consists of approximately 60 conserved amino acids [Shigyo et al., 2006]. Further N-terminal domains present in several AP2/ERF family TFs are an acidic domain and a serine rich region, both potentially involved in transcriptional activation [Riechmann and Meyerowitz, 1998].

A significant form of post-transcriptional regulation of AP2/ERF-domain TFs is phosphorylation. Pti-4 is phosphorylated through the receptor kinase Pto which increases its affinity to the *PR* gene promoters [Gu et al., 2000]. The rice TF OsEREBP1 also exhibits enhanced binding to the target GCC-box upon phosphorylation [Cheong et al., 2003].

1.4.2 The bHLH and bZIP family TFs

bHLH domain TFs are found both in animals and plants where they function in diverse processes. In plants bHLH proteins are involved in anthocyanin biosynthesis, phytochrome signalling and other plant specific processes [Buck and Atchley, 2003]. The functional motif includes a DNA-binding domain and a region comprised of two α-helices separated by a loop which confers homo- and heterodimerisation [Atchley and Fitch, 1997]. The binding motif of bHLH TFs is the E-box (G-box in plants); however, several binding motifs have been identified in animals, metazoan and fungi [Buck and Atchley, 2003]. The G-box promoter element has been shown to be jasmonate-responsive; however, bHLH TF binding sites are present in a multitude of non-jasmonate-responsive genes [Guedes-Correa et al., 2008, Buck and Atchley, 2003].

In flavanoid biosynthesis in maize the R/C1 protein, composed of a bHLH/ MYB TF heterodimer respectively coordinately regulates the anthocyanin structural genes [Mol et al., 1998]. Similar bHLH interactions have been found in *Petunia hybrida* and *Anthirinum majus*. In *A. thaliana* biosynthesis of the seed coat tannin is also regulated by bHLH/MYB proteins [Endt et al., 2002]. A jasmonate-responsive bHLH TF has been identified in *C. roseus*. CrMYC2 interacts with the G-box like element in the *STR* promoter which forms part of the JERE [Davies and Schwinn, 2003].

The bZIP domain TFs form a large multigene family conserved in all higher eukaryotes [Deppmann et al., 2006]. This family of TFs exhibit a highly conserved basic DNA-binding domain and a coiled coil leucine zipper domain [Hurst, 1995]. The promoter binding sites of bZIPs TFs are G-box related

sequences most of which harbour the ACGT core element [Davies and Schwinn, 2003]. bZIP TFs also interact with the as-1-type cis element present in the promoter of *PR* genes. Although the *LAP* gene and *A. thumefaciens* T-DNA exhibits a jasmonate-responsive as-1 type promoter element, bZIP binding to as-1 elements in the *PR* gene promoter occurs in pathogen-induced and salycilic acid dependent systemic acquired resistance [Schuetze et al., 2008].

The bZIP TF family functions in the transcriptional regulation of a variety of crucial cellular and developmental processes [Amoutzias et al., 2007]. The observed functional versatility can partly be attributed to dimerisation of bZIP TFs with other TF homo- or heteromonomers [Klemm et al., 1998]. Another advantage of dimerisation, which is conferred by the leucine zipper domain is the expansion of DNA-binding specificity and contact with novel DNA motifs [Cai et al., 2008]. As seen with AP2/ERF TFs phosphorylation is a common theme, regulating bZIP TF activity in plants [Schuetze et al., 2008].

1.4.3 Pathogen-responsive WRKY family TFs

The functional DNA-binding domain of the WRKY transcriptional regulators consist of a zinc-finger like motif and an N-terminal amino acid stretch [WRKYGQK] which is discriminatory for WRKY TFs [Rushton et al., 1996]. Only found in plants, WRKY TFs are further classified into three groups with group I containing two WRKY domains and group II only one WRKY domain. Group III consists of members of both groups I and II, which show a distinction within their zinc-finger like motif [Eulgem et al., 2000]. DNA-

binding of group I WRKY TFs is performed by the C-terminal domain [Hara et al., 2000]. Indeed sequence analysis showed a closer resemblance of the C-terminal domain of group I TF to the single WRKY domain of group II. The N-terminal domain of group I WRKYs might participate in DNA binding or the creation of a binding platform for protein interaction [Eulgem et al., 2000].

WRKY TFs recognise the W-box promoter element [(T)(T)TGAC(C/T)] as their target binding sequence which contains the TGAC core essential for WRKY interaction [Rushton et al., 1995]. WRKY TFs regulate pathogen elicited gene expression and W-boxes are present in promoters of genes associated with defence response [Rushton and Somssich, 1998]. The WRKY1 TF in parsley is rapidly induced and mediates defence responses to fungal pathogens [Eulgem et al., 1999].

In plant secondary metabolism a WRKY TF was recently shown to coordinately regulate the expression of berberine pathway genes through interaction with a W-box present in target promoters in cell cultures of *Coptis japonica* [Kato et al., 2007]. In accordance with pathogen induced accumulation of secondary metabolites, a csc from *Taxus chinensis* accumulates Taxol^{TM} in response to a fungal elicitor [Xu et al., 2004].

1.4.4 The MYB TF family

The MYB DNA binding domain of this class of transcriptional regulators is composed of a helix-helix-turn-helix motif. Most plant MYB TFs exhibit two repeats of the functional motif (R2R3 class) both of which are required for DNA interaction. However, the first MYB domain TF isolated, the C1 protein

in maize, harbours three motif repeats [Jin and Martin, 1999]. As MYB family proteins are classified by their DNA-binding motif structure, the MYB target binding site does not conform to a uniform consensus motif. Animal three-repeat and some plant two-repeat MYB TFs recognise the MYB binding site I: [T/CAACG/TGA/C/TA/C/T]; however, the majority of plant R2R3 MYB TFs will bind to the MYB binding site II: [TAACTAAC] [Romero et al., 1998].

Since the isolation of the C1 protein in maize, MYB TFs have been implicated in transcriptional regulation of plant secondary metabolism. Several R2R3 MYB TFs controlling anthocyanin biosynthesis in maize, ZmMYBPL, ZmMYB1 and ZmMYB38, *Petunia hybrida*, PhMYBAN2 and PhMYB2 and in *Anthirinum majus*, AmMYB305, have been identified [Jin and Martin, 1999]. In tobacco a jasmonate-responsive MYB TF has been shown to regulate phenylpropanoid biosynthesis [Galis et al., 2006]. Furthermore MYB TFs in accordance with bHLH proteins regulate flavanoid biosynthesis in maize [Mol et al., 1998].

1.5 The biology and chemistry of Taxol™

Interest in Taxol™ can be largely accredited to the successful clinical trials which started in 1984. However, already in Phase II of the trials supply constraints were apparent, which were later overcome by the semisynthesis of Taxol™ [Goodman and Walsh, 2001]. Today Taxol™ (Bristol-Mayer-Squibb) and decetaxel (Taxotere, Aventis) are in clinical use. These two taxenes have been shown to be successful in the treatment of ovarian, breast and non-small

lung cancer in mono and combinational therapy. Effects of taxenes have also been shown in prostate, head and neck, gastric, cervical and germ-cell cancer and the HIV-related Kaposi's sarcoma [Crown and O'Leary, 2000].

1.5.1 The history of Taxol™

The Cancer Chemotherapy National Service Center was created by the National Cancer Institute in 1955 as a drug research and development program with collaborators in academics and industry. In order to explore plant natural products as anticancer drugs, the CCNSC carried out a screen for drugs with *in-vivo* antitumor activity using the leukaemia mouse lines L1210 and P388. The screen was led by the USDA and involved 114,000 plant derived extracts [Cragg, 1998]. In 1962 plant samples of stem, fruit, leaves, twig and bark of *Taxus brevifolia* were submitted for screening [Goodman and Walsh, 2001]. The bark extract from *Taxus brevifolia* showed activity in the leukaemia models and in a subsequent KB cytoactivity assay. The sample was recollected in order to isolate the active compound, which succeeded in 1967 and was named Taxol™. In 1971 the chemical structure of Taxol™ was resolved by X-ray analysis and ^{1}H-NMR as a tetracyclic, highly oxygenated diterpene with a ester side chain at carbon atom 13 (C-13) [Cragg, 1998]. Further scientific interest was caused by Dr. Susan Horwitz who first showed insights into the Taxol™ mechanism of action as a tubulin promoting agent [Horowitz et al., 1979].

Although the extract from *Taxus brevifolia* showed activity in the initial bioassays, it was only one of numerous entries and was therefore shown but

moderate interest. However, the extract containing Taxol™ made its breakthrough when tested against the B16 melanoma model and was subsequently put forward for clinical trials [Cragg, 1998].

En route to the availability of Taxol™ as an anticancer drug in clinical use numerous challenges had to be overcome, the most apparent being supply constraints. The initial yield of only 0.01% of Taxol™ from the bark of *Taxus brevifolia* highlights the difficulty to obtain sufficient amounts necessary for clinical trials. This equates to one kg of Taxol™ to about 3000 yew trees. This was compounded by the high dose of Taxol™ requirement in clinical use [Goodman and Walsh, 2001].

1.5.2 Taxol™ biosynthetic pathway

The acyclic precursor for Taxol™ biosynthesis, GGPP (C-20) [Fig.: 1.4], is supplied by the GGPPS which couples IPP (C-5) with FPP (C-15). Three IPP units, which derive from the plastidial MEP pathway, and DMPP, the IPP isomer, are required for GGPP synthesis. The enzyme constitutes the branch-point progenitor for a variety of diterpenoids and tetraterpenoids, several of which are relevant to the process of photosynthesis [Hefner et al., 1998] [Walker and Croteau, 2001]. The cyclisation of GGPP to taxa-4(5),11(12)-diene by the diterpene cyclase TASY, is the committed step which forms the taxane skeleton [Wildung and Croteau, 1996]. The cytochrome P450-mediated hydroxylation and double bond migration by T5αH yields taxa-4(5),11(12)-diene-5-α-ol. At this point a branch in the metabolic pathway occurs. The enzyme TαH catalyses the conversion of taxa-4(20),11(12)-dien-5-α-ol to taxa-4(20),11(12)-

diene 5-α, 13-α-diol [Jennewein et al., 2001]. The alternative branch in the pathway implements TDAT to form taxa 4(20),11(12)-diene-5-α-yl acetate from taxa-4(20),11(12)-dien-5-α-ol [Walker et al., 1999], which is then further converted to taxa-4(20),11(12) diene 5-α-acetoxy-10-β-ol by T10βH [Schoendorf et al., 2001]. The steps leading from the acetate or the diol intermediates

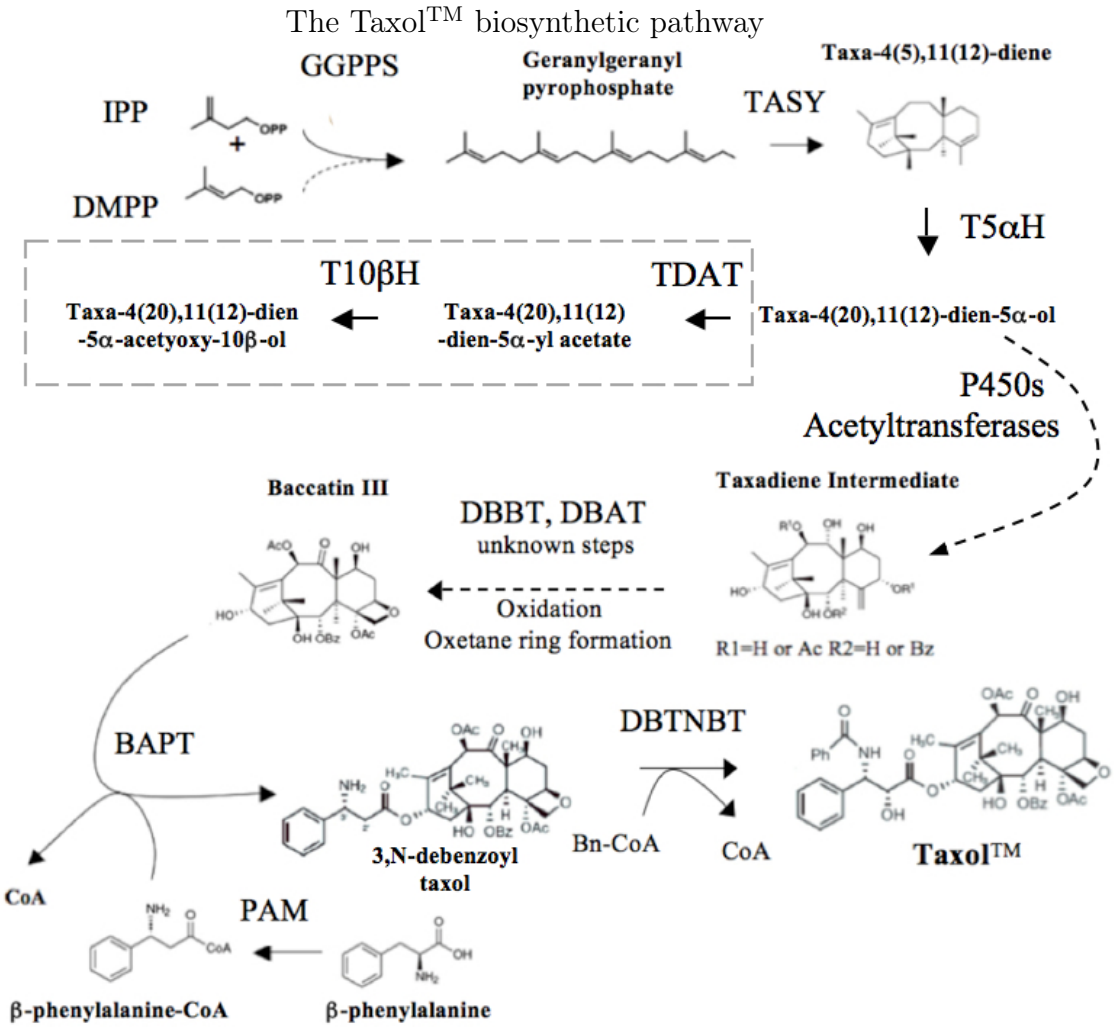

Figure 1.4: Taxol™ biosynthetic pathway. The pathway branch is shown in the dashed box.

to functionalized taxanes are unknown. Highly modified taxane production occurs via taxane DBBT [Walker and Croteau, 2000a] to produce 10-DAB, which is then converted to baccatin III by DBAT [Walker and Croteau, 2000b]. The transferase BAPT catalyses the coupling of the β-phenylalanine side chain (derived from phenylalanine via PAM [Jennewein et al., 2004]) to baccatin III to produce 3′N-debenzoyl taxol [Walker et al., 2002]. The final step yielding Taxol™ is catalysed by DBTNBT which couples a benzoyl group to 3′N-debenzoyl taxol [Walker et al., 2002]. BAPT and DBTNBT have recently been identified as putative rate limiting steps of the Taxol™ pathway [Nims et al., 2006].

1.5.3 Taxol™ bioactivity

With the elucidation of the chemical structure of Taxol™ its cellular effect in interfering in microtubule function became apparent [Horowitz et al., 1979]. Microtubules are cellular structural components and are essential for proper cellular movement, cell shape integrity and cytokinesis. The building blocks of microtubules, which are highly dynamic protein polymers, are $\alpha\beta$-tubulin heterodimers. Polymerisation and depolymerisation occurs in a tightly regulated reversible hydrolysis where GTP binds the β-tubulin subunit and is converted to GDP and orthophosphate [Hyams and Lloyd, 1994].

Several structural features of Taxol™, including the N-benzoyl phenylalanine C-13 side chain, the acetate at C-4, the benzoate group at C-2, the oxetane ring and the cup shaped taxane core itself are important elements contributing to the activity and tubulin binding [Croteau et al., 2006]. Bind-

ing of Taxol™ occurs on the inner surface with one molecule per heterodimer along the length of the microtubule [Yeung et al., 1999]. Taxol™ can associate to either GTP- or GDP- bound tubulin resulting in a conformational change of the heterodimer thus stabilising the microtubules [Arnal and Wade, 1995]. The primary effect is the prevention of dynamic changes of microtubules of the mitotic spindle, which leads to arrest in G2/M phase and results in failure of chromosome segregation and leads ultimately to apoptosis [Blagosklonny and Fojo, 1999]. The cell cycle arrest and inhibition of proliferation is of particular advantage in cancer cells, which exhibit a fast growth rate [Cho et al., 2008].

Several downstream effects of mitotic arrest have been described including hyperphosphorylation of numerous proteins including Bcl2, a protein with antiapoptotic function, thus initiating the apoptotic program. Another hyperphosphorylated protein in Taxol™ cytotoxicity is the cell cycle checkpoint TF p53. This tumor suppressor protein can induce growth arrest and apoptosis. Other effects induced by Taxol™ and mitotic arrest are the induction of cytokines and expression of early response genes, including TFs with tumor suppressor activities [Blagosklonny and Fojo, 1999, Blagosklonny, 2002].

In 1984 phase I of the Taxol™ clinical trials started followed by phase II in 1985. Although supply constraints limited the trials especially in phase II to a smaller number of participants, Taxol™ administration showed clear clinical responses in breast and ovarian cancer [Hoff, 1998,Holmes et al., 1991].

1.5.4 Supply and synthesis of Taxol™

Taxol™ extraction was confined to the same method and source (bark) from

1962 until 1993. From 1978 onwards Hauser Chemical Research of Boulder was contracted by the National Cancer Institute to manage the bark extracts necessary for clinical trials. However, it was clear that a more sustainable method had to be found as removing the bark kills the slow growing yew tree [Goodman and Walsh, 2001].

A method for semi-synthesis of Taxol^{TM} was published in 1981. This process used the precursor 10-deacetylbaccatin III, which can be readily extracted from harvested needles of yew. The advantage of this method is that it used a renewable plant source and doesn't interfere with the tree's growth [Denis et al., 1988]. Shortly afterwards a similar method was put forward using the same precursor where the Taxol^{TM} yield was increased by 80% (Holton, 1990). This process was adopted by the manufacturer Bristol-Myers Squibb, who trademarked the name of the drug Taxol^{TM} [Morrissey, 2003].

1.6 The model plant *Taxus cuspidata*

The plant material utilised in this work is a procambium csc from *Taxus cuspidata* (Siebold and Zucc.). A homogeneous cell layer of procambium cells was separated from a young twig [Fig.: 1.5] and cultured on solid medium [Lee et al., 2010]. The procambium is a primary meristem derived from the apical meristem and gives rise to primary xylem and phloem [Donner et al., 2010]. It also produces the vascular cambium from which secondary xylem and phloem derive. Differentiation into the conductive tissue occurs during the process of xylogenesis [Fukuda, 1996].

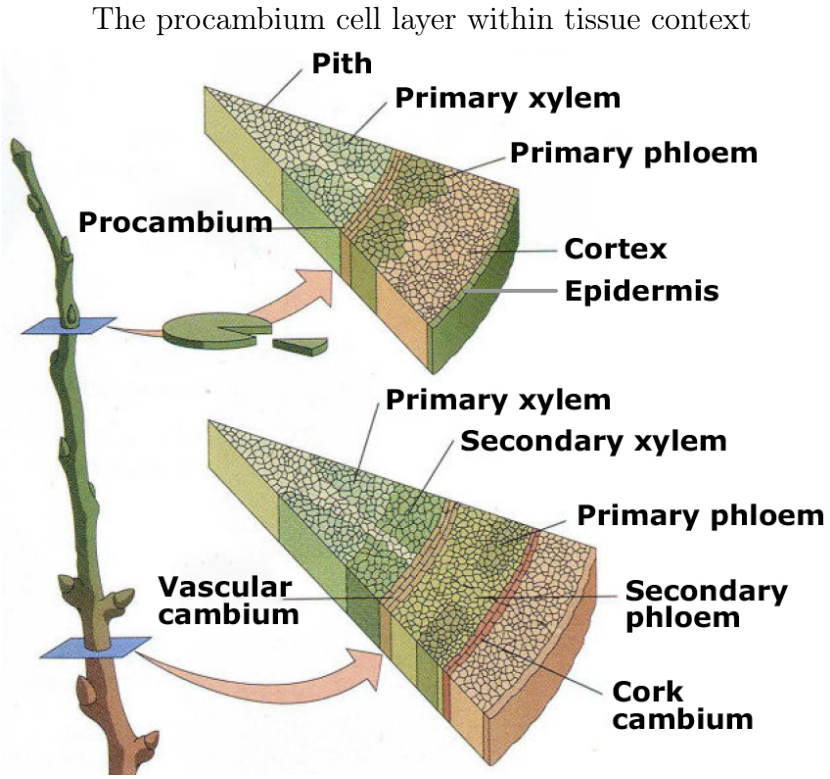

Figure 1.5: Cross section of twig showing the distinct cell layers including the procambium. Source: [Lee et al., 2010]

Cultured plant cells acquire a state which enables the regeneration of whole plants [Johri and Bhojwani, 1965]. The process of callus formation was described as de-differentiation of somatic cells which is accompanied by the withdrawal from a given differentiated cell morphology [Grafi and Avivi, 2004]. This was recently redefined as the acquisition of a pericycle cell-like state [Sugimoto et al., 2010]. Yet differential gene expression patterns were observed among callus forming cells suggesting a degree of heterogeny among callus cells [Gordon et al., 2007]. This heterogeny was observed in *Taxus* csc which may contribute to the observed variability of Taxol^{TM} accumulation from cell lines induced from somatic cell types [Naill and Roberts, 2005a, Ketchum

et al., 1999]. In contrast the employed csc was induced from procambium cells which had never entered a differentiated cell state [Lee et al., 2010]. Associated positive traits include stable and increased Taxol^{TM} accumulation and drastically reduced cell aggregation. An especially larger aggregate size can hamper nutrient and oxygen supply thus altering growth behaviour of csc [Hulst et al., 1989] [Pepin et al., 1999]. Furthermore cell aggregation can interfere with secondary metabolite accumulation. In *Salvia officinalis* and *Daucus carota* csc large cell aggregates negatively influenced production of secondary metabolites [Bolta et al., 2003] [Madhusudhan and Ravishankar, 1996]. However, large aggregate size favours the accumulation of anthocyanins in *Fragaria ananassa* csc [Edahiro and Seki, 2006]. Although it has been shown that large aggregate size in a *Taxus cuspidata* csc increases Taxol^{TM} accumulation by 5 fold to 4.9 mg/L in 50 ml culture [Xu et al., 1998], cell aggregation and associated features such as limited oxygen supply hamper stable large-scale growth in a bioreactor environment [Roberts, 2007].

Working with *Taxus cuspidata*

Working with the *Taxus cuspidata* procambium csc is impeded by characteristics such as a high phenolic content [8 mg/ g dry weight] as compared to the somatic csc [0.7 mg/ g dry weight] [Lee, personal conversation] and with 10,850 Mb (1C), a large genome [Murray, 1998]. *Taxus cuspidata* is a diploid plant with 24 chromosomes [Murray, 1998].

Plant metabolites such as phenolic compounds and terpenes interfere during the extraction of nucleic acids and can react irreversibly with proteins and

nucleic acids [Katterman and Shattuck, 1983]. Residual phenolic compounds present in extracted nucleic acids render them unsuitable for enzymatic restriction and amplification by *Taq* DNA polymerase [Porebski et al., 1997, Michiels et al., 2003].

.

1.6.1 *Taxus* as plant model

Since the early 1970s *Taxus* species were utilised as model plants on the background of Taxol^{TM} research. Initial studies concentrated on the isolation and structural analysis of Taxol^{TM} [Wani et al., 1971], Taxol mode of action [Horowitz et al., 1979, Fuchs and Johnson, 1978, Jacrot et al., 1983, Riondel et al., 1986], the isolation of further Taxenes [McLaughlin et al., 1981, Kingston et al., 1982, Huang et al., 1986] and Taxol^{TM} content analysis in different species of *Taxus* and tissue types [Vidensek et al., 1990, Witherup et al., 1990].

The structural characterisation of Taxol^{TM} as a diterpenoid increased scientific interest in this class of taxenes, resulting in more than 550 identified and isolated compounds from *Taxus* species. Chemical surveys in *Taxus cuspidata* shown a total of 152 diterpenoids and 41 non-taxene compounds among which are sesquiterpene, ecdysteroids, steroids, lignans, flavonoids and others [Wang et al., 2010].

Due to attempts to utilise the plant cell culture platform for large-scale production of Taxol^{TM}, numerous reports have been published concentrating on optimisation of cell type, medium composition, growth condition and elic-

itor studies for *Taxus* csc [Ellis et al., 1996, Wang et al., 2001, Wu and Lin, 2003].

1.6.2 Botany of *Taxus*

Taxus cuspidata (Siebold and Zucc.) or the Japanese yew is a small coniferous tree up to 16 m tall and grows in a pyramidal form with spreading or ascending branches [Hartzell, 1991]. The shoots of the tree are red-brown with spirally arranged, linear dark green leaves of 15-25 x 2-3 mm [Ohwi, 1965]. The Japanese yew is in the Taxaceae family. The genus *Taxus*, which is largely confined to the middle latitudes of the northern hemisphere, belongs to the taxonomic order Taxales [Fig.: 1.6]. Plants in this order bear an aril, a fleshy envelope, which surrounds the seeds, the distinctive feature distancing them from the Coniferales. Yew trees are extremely slow growing with germination taking up to two years and for a yew to reach its mature height about 200 years [Suffness, 1995].

Taxus cuspidata is native to Japan; China (Heilongjiang, Jilin, Liaoning, Shaanxi); the Republic of Korea (South), the Democratic People's Republic of Korea (North) and the Russian Federation (Kuril, Primorye, Sakhalin) [IUC, 2010]. *Taxus* species can be found from almost sea level to an altitude of 3000 m along the entire latitude [Fig.: 1.7] [de Laubenfels, 1988]. Within the genus *Taxus* are nine species, which are namely: *T. baccata* (L.) or the European yew; *T. brevifolia* (Nutt.) or the Pacific or the Western yew, *T. canadensis* (Marshall) or the Canadian yew; *T. chinensis* (Pilg. Rehder) or the Chinese yew; *T. floridana* (Nutt. ex Chapman) or the Florida yew; *T.*

globosa (Schltdl.) or the Mexican yew; *T. sumatrana* (Miq. de Laub.) or the Sumatran yew, *T. wallichiana* (Zucc.) or the Himalayan yew and *T. cuspidata* (Siebold and Zucc.) [Hartzell, 1991]. The occurrence of *Taxus* is dated to 200 million years ago, proven by a fossil record of *Paleotaxus rediviva*, whose structure is identical to the modern yew. In 1998 *Taxus cuspidata* was assessed by the Conifer-Specialist-Group with a lower risk status in the IUCN Red List of threatened species [IUC, 2010].

Taxus cuspidata

Figure 1.6: Plate from Flora Japonica, Sectio Prima, showing *Taxus cuspidata* (Siebold and Zucc.), published in 1870 [Illustration 128]. Source: [von Siebold and Zuccarini, 1870].

Taxus habitats

Figure 1.7: Map showing the natural habitat of *Taxus* species labelled in red. Source: [URL: http://www.conifers.org/ta/ta/index.htm; The Conifer Database.]

Taxus cuspidata is distantly related to other gymnosperm organisms used as models such as the pine and *Picea* species. The latter are found among the Pinophyta within the Pinacea family which belongs to the subclass of the pinidae. In contrast *Taxus cuspidata* within the Taxaceae is found within the subclass of the Taxidae [Fig.: 1.8]. The divergence of the Taxidae and Pinidae from the Corditidae is estimated at about 300 MYA [Sitte et al., 1991]. The well characterised angiosperm plant model *A. thaliana* is found among eudicots within the Brassicales. Other frequently used and well characterised plant models are *Oryza* and *Zea* species found within the monocots. The divergence time of the eudicots and monocots within the angiosperms is estimated at 131-200 MYA [Nam et al., 2003].

Divergance of gymnosperm and angiosperm model organisms

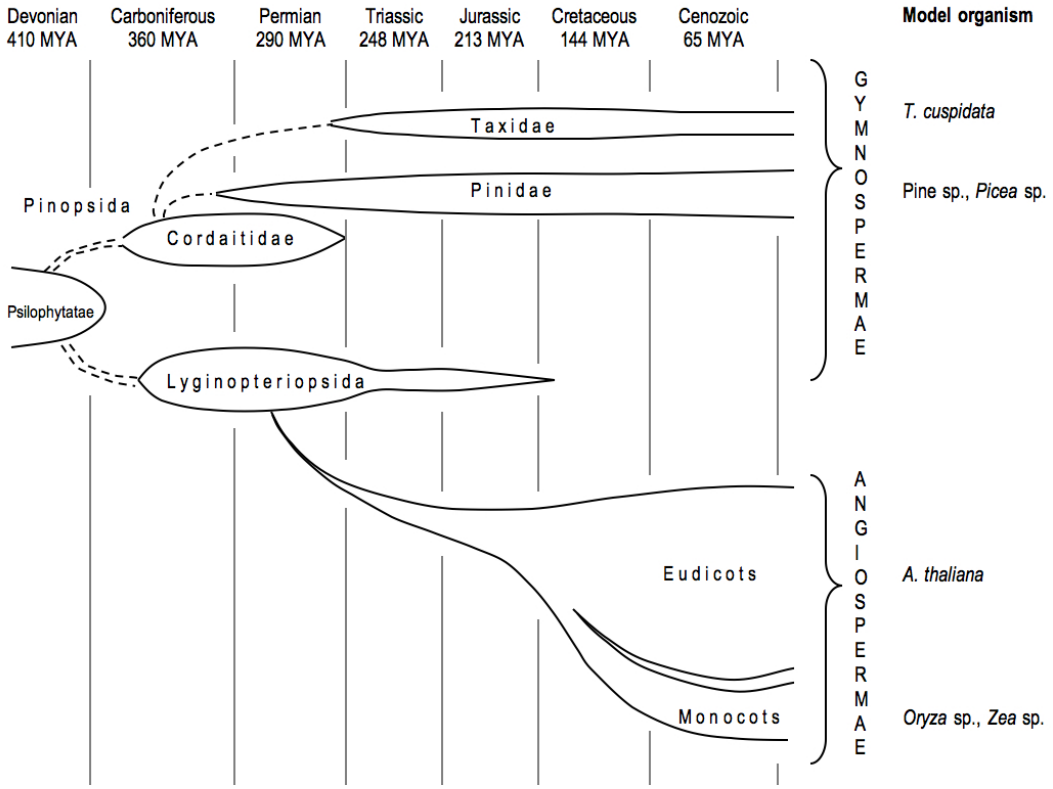

Figure 1.8: The diagram shows the evolution of angiosperms and gymnosperms from the Psilophytatae, the Progymnospermae, over time and highlights the divergence and relationship of plant model organisms [adapted from [Sitte et al., 1991]

Ethnobotany of *Taxus*

The yew tree was treasured because of its hard durable wood, feared due to its toxic foliage and worshiped by several cultures throughout history. The name *Taxus* is derived from the Greek word for yew, toxos, which is thought to originate from the combination of the Greek words toxicon and toxon meaning

poison and bow, respectively. This sums up the valuable characteristics of the yew tree known by the ancient Egyptian, Greek, Roman, Celtic and Asian cultures. Each culture gave the yew tree its unique connotation. However, it was always a sacred tree associated with long life and the underworld [Suffness, 1995].

The yew tree was used in traditional medicine to treat tuberculosis and epilepsy. A tea from yew was used to initiate abortions in medieval times and an extract, brewed from yew foliage was used as poison. The yew tree gained entry into modern medicine in the late 1970s with the discovery of Taxol^{TM} [Suffness, 1995].

1.7 Project aims

The non-model organism *Taxus cuspidata* produces the potent anti-cancer drug Taxol^{TM} which makes it attractive for molecular studies concerned with the regulation of the Taxol^{TM} biosynthesis pathway. In this project molecular resources and methodology were established necessary for the identification and characterisation of transcription factors which govern Taxol^{TM} biosynthesis. A further aim was the characterisation of the employed csc which was initiated from *T. cuspidata* procambium tissue.

Specific aims of the project were as follows:

• Cloning of Taxol^{TM} biosynthetic gene promoters and their In-silico analysis.

• Establishment of a transient transformation assay using *T. cuspidata*

csc for functional characterisation of the isolated Taxol^{TM} biosynthetic gene promoters.

• Sequencing of the *T. cuspidata* transcriptome.

• Identification of jasmonate-responsive transcription factors from *T. cuspidata*.

• Characterisation of Taxol^{TM} biosynthetic gene expression in the procambium and somatic csc.

• Characterisation of the employed procambium csc by identification of marker genes.

Chapter 2

Materials and Methods

2.1 Plant material

Taxus cuspidata cell suspension cultures

Taxus cuspidata callus was obtained from the biotechnology company Unhwa [Lee et al., 2010]. Both the homogenous procambium [CMC] and the heterogenous somatic [DDC] callus were propagated and maintained at 20°C in the dark on petri-dishes containing 35 ml callus growth medium [Table.: 2.1] supplemented with 0.5 mg/L gibberellic acid∗ and 1% sucrose∗, 0.4% gelrite and 0.01% activated charcoal. The CMC callus was sub-cultured every 14 days and the DDC every 21 days. A cell suspension culture [csc] from *T. cuspidata* CMC callus was established using 2.5 g of 10 day old callus suspended in 40 ml of liquid growth medium supplemented with 1 mg/L gibberellic acid∗ and 2% sucrose∗. The medium was adjusted to pH 5.5 and autoclaved. Subculturing of the csc 1/10 (v/v) was carried out every 10-14 days. [Asterisks

Taxus medium salts, amino acids and other compounds	
Compound	Final conc. mg/L
Potassium nitrate	1011.1
Magnesium sulphate heptahydrate	121.56
Manganous sulfate tetrahydrate	10
Zinc sulphate heptahydrate	2
Cupric sulfate pentahydrate	0.025
Calcium chloride monohydrate	113.23
Potassium iodide	0.75
Cobalt(II) chloride hexahydrate	0.025
Monobasic sodium phosphate, monohydrate	130.44
Boric acid	3
Sodium molybdate dihydrate	0.25
Iron sodium ethylenediaminotetraacetate	36.7
Myo-inositol	100
Thiamine-hydrocloride	10
Nicotinic acid	1
Pyridoxine hydrochloride	1
L-Aspartic acid	133
L-Arginine	175
Glycine	75
L-Proline	115
L-Ascorbic acid∗	100
Citric acid∗	150
Picloram	1

Table 2.1: *T. cuspidata* medium composition

indicate compounds which were dissolved in 100 ml of dH2O, filter sterilised and added after autoclaving.]

Tobacco plants

Nicotiana benthamiana plants utilised in transient agroinfiltration assays were grown under long day conditions [16 h light, 8 h dark, light intensity 1000 μmol m^{-2} sec^{-1}] at 25° C.

Methyljamonate treatment of *Nicotiana benthamiana* plants

A 100 mM stock solution of MeJA in ethanol was prepared and stored at -20°C. For treatment of *N. benthamiana* plants the stock was further diluted to 200 μM and equally sprayed over the leaf areas previously infiltrated with *Agrobacterium*.

2.2 Plant transformation methods

Particle bombardment of *Taxus cuspidata* csc

In order to transform the utilised plant material by particle bombardment the DNA construct designed for transformation was coated onto the microcarriers (gold particles) [Bio-Rad, USA]. A suspension was prepared containing 50 mg gold microcarriers in 1 ml of sterile H_2O by vigorous vortexing for 1-2 min. The suspension was subsequently centrifuged at 13,000 rpm for 1 min and the supernatant was carefully discarded. The microcarriers were washed using 100% ethanol and subsequently centrifuged three times. To remove the ethanol a wash using sterile H_2O was performed. After discarding the supernatant, the prepared microcarriers were resuspended in 1 ml of sterile H_2O. 5 μg of plasmid DNA were added to an aliquot of 25 μl of microcarrier suspension and briefly vortexed. The DNA was precipitated onto the microcarriers using 25 μl of 2.5 M $CaCl_2$. The suspension was mixed gently by pipetting and 10 μl of 0.1 M spermidine was added. Both $CaCl_2$ and spermidine are essential for good DNA

precipitation and exclusion of spermidine results in decreased transformation efficiency [Zuraida et al., 2010]. Precipitation of DNA by high $CaCl_2$ concentration is caused by competition of salt ions with H_2O molecules [Edsall and Wyman, 1958]. The multivalent cation spermidine, binds to, and condenses the DNA in aqueous solutions [Wan and Wilkins, 1993, Pelta et al., 1996]. After gentle mixing the suspension was gently vortexed for 2 min and incubated on ice for 30 min. The pelleted microcarriers were subsequently washed using 200 μl of 100% ethanol and after further centrifugation resuspended in 100 μl of 100% ethanol.

Particle bombarment was performed with various rupture disc pressures including 650 psi, 1100 psi and 1350 psi. The vacuum pressure at the time of bombardment was 28 mmHg. 1.0 μm and 1.6 μm microcarriers were utilised at a target distance of 3 cm.

Isolation and electroporation of *Taxus cuspidata* protoplasts

The enzymes used for protoplast isolation (1 % [w/v] Cellulase-*Trichoderma*, 0.1% [w/v] Pectolyase Y-23 [MP Biochemicals, UK] and 0.2% [w/v] Driselase [Sigma-Aldrich, USA]) were dissolved in 5 ml TEX buffer [500 mg/L MES (2-(N-Morpholine)-ethane sulfonic acid), 750 mg/L calcium chloride dihydrate, 250 mg/L ammonium nitrate, 0.5 M mannitol, pH 5.7 with KOH] by vigorous shaking and left at room temperature for 30 min. The enzyme mix was then centrifuged at 10,000 g for 10 min to remove insoluble particles and the supernatant was used for isolation of protoplasts. Protoplasts from 2 ml settled cell volume of procambium csc from *T. cuspidata* were generated by enzymatic

digestion in 7 ml of enzyme mix overnight in the dark. The digestion was performed in a petri-dish to ensure sterile conditions. Prior to protoplast purification the petri-dish was shaken gently to liberate the maximum number of protoplasts. The protoplast suspension was then gently transferred on top of 10 ml of 16% sucrose solution and centrifuged for 15 min at 100 g without break. The viable protoplasts were harvested from sucrose-enzyme solution interphase and washed in 10 ml electroporation buffer [0.4 M sucrose, 2.4 g/L hepes, 6 g/L KCl, 600 mg/L calcium chloride dihydrate pH 7.2 with KOH]. The protoplasts were then passed through a 100 μm nylon filter which previously had been wetted with electroporation buffer.

After a further centrifugation for 15 min at 100 g without break the protoplasts were resuspended in an appropriate volume of electroporation buffer in order to obtain 3 X 10^6 protoplasts/ ml [according to the thickness of the isolated protoplast layer].

500 μl of the isolated and purified protoplast suspension were transferred into a plastic cuvette, and mixed with 100 μl of electroporation buffer containing 50 μg of DNA. The DNA/ cell mixture was incubated for 5 min at room temperature. The electroporation is performed with 910 μF and the optimum voltage for *T. cuspidata* procambium protoplasts of 170 V by lowering the electrodes into the plastic cuvette. Electroporation was performed with a home-built device [by Prof. J. Denecke, Leeds] with an electrode distance of 3.5 mm. After electroporation the protoplasts were left stationary for roughly 15 min and then diluted five times in TEX buffer. Gene expression of the transformed protoplasts was analysed after an overnight incubation in the

dark at room temperature.

Agrobacterium infiltration of *Nicotiana benthamiana* plants

Single colonies of the transformed *Agrobacterium* strain were grown in a 5 ml volume of LB-medium at 28°C for 1-3 days and centrifuged at 2,500 rpm for 15 min at 16°C. The pellet was subsequently resuspended in 1 ml of infiltration buffer [0.01 M magnesium chloride, 0.01 M MES, pH 5.6 KOH, 15 μM acetosyringone] and the OD_{600} was determined of a 1 in 100 dilution of the resuspension. The *Agrobacterium* resuspension was diluted to the appropriate OD. Immediately prior to infiltration a syringe tip was used to make a fine incision into the lower surface of a fully expanded *N. benthaminana* leaf at the side of infiltration. A volume of 100 μl was infiltrated into the leaf using a 1 ml syringe by applying gentle counter-pressure to the upper leaf surface. In this way 4 - 6 infiltrations were performed on one leaf. Jasmonate treatment was performed 12 hours after infiltration [Yang et al., 2000, Voinnet et al., 2003].

2.3 Reporter assays

Fluorometric GUS activity assay

GUS activity was measured by monitoring cleavage of the β-glucuronidase substrate MUG which yields 4-MU and sugar glucuronic acid [Jefferson, 1989, Gallagher, 1992]. Roughly 50 mg of plant tissue were ground in liquid nitrogen

and added to 500 μl of GUS extraction buffer [50 mM Sodium phosphate pH 7.0, 10 mM β-mercaptoethanol, 10 mM EDTA, 0.1% sarcosyl, 0.1 % [v/v] triton X-100]. The samples were stored on ice and centrifuged at 12,000 rpm for 10 min at 4°C. Subsequently 100 μl of the resulting protein extract were added to 900 μl of GUS assay buffer [GUS extraction buffer containing 2 mM MUG]. The samples were then incubated at 37°C. The reaction was stopped after 1 hour by the addition of 200 μl of GUS stop solution [0.2 M Sodium carbonate] followed by vortexing of the sample. Using excitation of 365 nm and measuring emission at 455 nm, the amount of 4-MU produced from the substrate 4-methylumbelliferyl β-D-glucuronide was quantified.

To normalise the GUS activity to the imput protein concentration of the used samples a Bradford assay was performed. 5 μl of protein extract used to measure GUS activity was added to 200 μl Bradford reagent and 795 μl H2O. The resulting protein concentration was used to express the GUS activity as pmol MU/ min/ mg protein.

α-Amylase activity assay

α-Amylase [1,4-α-D-glucan glucanohydrolase] hydrolyses blocked p-nitrophenyl maltoheptaoside into blocked maltosaccharide and p-nitrophenyl malto-saccharide. The latter is further hydrolysed by the action of a thermostable α-glucosidase to yield p-nitrophenol and glucose [Nater et al., 2005].

Transformed protoplasts utilised to measure the α-amylase activity were centrifuged at 13,000 rpm for 4 min and the pellet was resuspended in 500 μl of extraction buffer [50 mM malic acid, 50 mM sodium chloride, 40 mM calcium

chloride, 0.02 % sodium azide, 0.02 % BSA, pH 5.4]. Following sonication [Ultrawave Ltd, UK] for 5 sec at amplitude 10 the samples were centrifuged at 13,000 rpm for 4 min. 30 μl of the supernatant were added to 30 μl substrate [Kit, Megazyme, Australia, containing: blocked *p*-nitrophenyl maltoheptaoside 5.45 mg/ml, α-glucosidase 100 U at pH 5.2]. The reaction was started by incubating the sample at 40°C and stopped after exactly 30 min by the addition of 150 μl stopping solution [1% (w/v) Trizma base, pH 11]. The generated *p*-nitrophenol absorbs light at a wavelength of 405 nm.

2.4 DNA manipulation and analysis

Extraction of *Taxus cuspidata* genomic DNA

About 300 mg of the *T. cuspidata* csc was collected and the culture medium was removed by placing the cells on top of a fine nylon mesh above a stack of tissue paper. The plant material was subsequently ground with 5 μl of 1% (v/v) 2-mercaptoethanol. After adding 300 μl of extraction buffer [250 mM NaCl, 25 mM EDTA, 0.5% SDS, 200 mM Tris-HCl pH 8.0] the homogenate was incubated for 1 hour at room temperature. Freshly prepared PVP [MW 10,000] solution [6% of the final volume] and half a volume of 7.5 M ammonium acetate were added separately. The homogenate was then incubated for 30 min on ice and centrifuged at 4°C at 10,000 g. The supernatant was carefully removed and transferred to a fresh tube to which an equal volume of ice cold isopropanol was added. Following an incubation for 30 min at -20°C to precipitate the DNA, the tube was centrifuged at 4°C for 30 min at 10,000 g. The supernatant was

discarded and the pellet was allowed to dry. The DNA was then resuspended in 500 μl TE buffer [10 mM Tris-HCl pH 8.0, 0.1 mM EDTA pH 8.0]. Two μl of RNase [1 mg/ml] were added to the solution and incubated at 37°C for 15 min. To remove the RNase and plant pigments, one volume chloroform-isoamyl alcohol (24:1) was added and after several inversions centrifuged at 10,000 g for 5 min at 4°C. The aqueous fraction was removed and transferred to a new tube. The chloroform extraction was repeated once more. The aqueous fraction was then transferred to a fresh tube and 1 volume of ice cold isopropanol was added and incubated for 30 min at -20°C. After centrifugation at 10,000 g for 30 min, the pellet was washed with 1 ml of 80% ethanol and air dried under sterile conditions. The pelleted DNA was subsequently redissolved in 30 μl TE buffer [Kim et al., 1997].

Cloning of *TASY::GUS* fusion construct

For the bombardment of *T. cuspidata* csc three plasmids were cloned [Fig.: 2.1] to assess the jasmonate inducibility of the isolated *TASY* promoter fragment. A fragment from the pJIT-166int vector [Fig.: 2.2 A] [Guerineau et al., 1992] harbouring the *GUS* reporter gene under the control of the constitutive *CaM V 35S* promoter was cloned into the *Sac*I - *Xho*I site of the pGreenII 0049 vector [Fig.: 2.2 B] [Hellens et al., 2000]. A *CaM V 35S* promoterless fragment from the pJIT-166int vector was also cloned into the *Hind*III - *Xho*I sites of the pGreenII 0049 vector. The isolated *TASY* promoter was cloned as *Hind*III - *Sal*I fragment into the latter.

Name	Sequence [5']
TASY hind-fwd:	GGCAAGCTTGTCCACTATCTACTTTGAAAATACCTCT
TASY sal-re:	ATAGTCGACTTCTGCAGAGAGGCAGGGGAAC

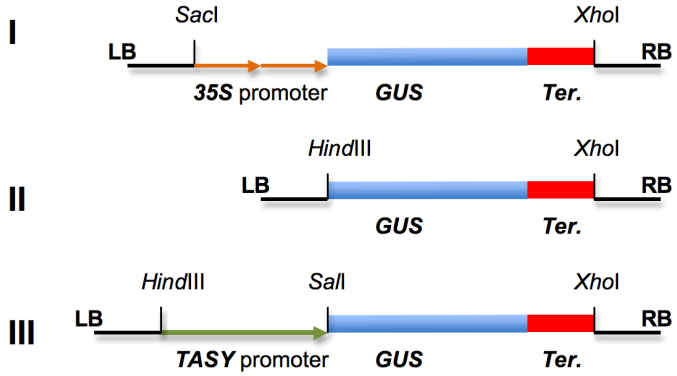

Figure 2.1: Cloned *GUS* reporter constructs; [I]: The *GUS* gene under the control of a double *CaM V 35S* promoter; [II]: A promoterless control construct; [III]: The *GUS* gene under the control of the *TASY* promoter.

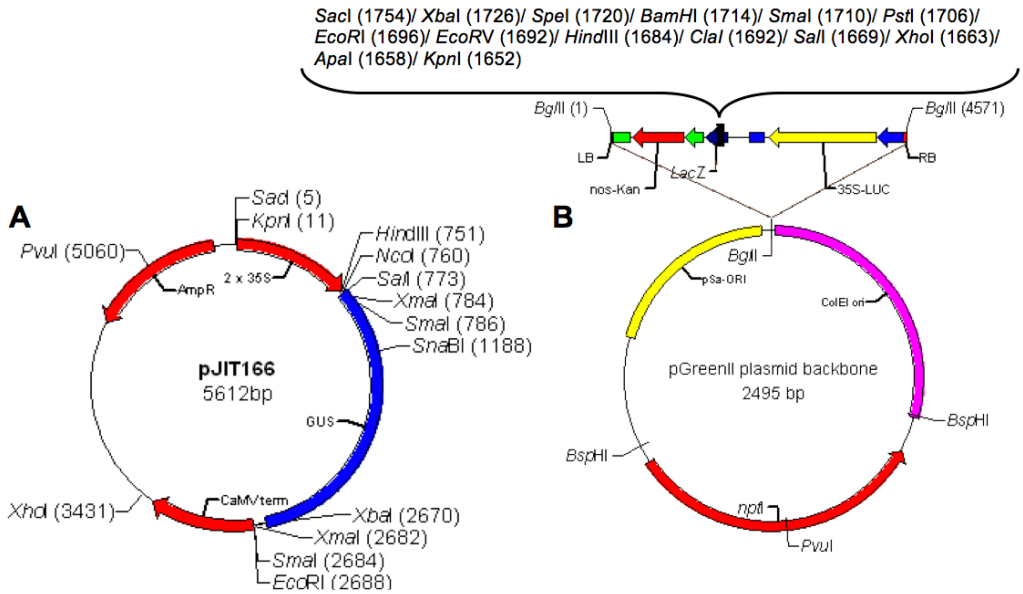

Figure 2.2: [A]: The pJIT166 vector harbouring the *GUS* reporter gene; [B]: The pGreen0049 vectors which has the left [LB] and right [RB] T-DNA integration border.

Gateway® BP and LR reactions

The Gateway® cloning manual can be found at: www.tcd.ie/Genetics/staff/Noel.Murphy/recombinant%20dna%20ge4021/gatewayman.pdf. The Gateway® cloning [Invitrogen, USA] technology involving the BP and LR recombination reactions was performed according to the manufacturer's instructions. This includes the PCR amplification of a desired DNA fragment using gene specific primers which have additional Att-overhangs. These support the site-specific recombination of the PCR fragment into the pENTRY^{TM} vector during the BP reaction. Following transformation of 1 μl of the recombined DNA from the BP reaction into *E. coli* strain XL1-blue^{TM}, the plasmid DNA was extracted and utilised in the LR reaction in which the DNA fragment of interest, now within the pENTRY vector, is transferred by site-specific recombination into an appropriate destination vector. Subsequent transformation into *E. coli* and sequencing of the extracted plasmid DNA was performed.

Gateway® overhangs:	
Name	Sequence [5']
Attb1:	GGGGACAAGTTTGTACAAAAAAGCAGGCTTC
Attb2:	GGGGACCACTTTGTACAAGAAAGCTGGGTG
Specific primers:	
*DBTNBT*fwd:	ATGGGGAACTGGAAGTGGAT
*DBTNBT*re:	GACTGGATCAAAGATGAAACGAT
*TASY*fwd:	AACTCGCAATAGCTAGGACATCTT
*TASY*re:	GCAGAGAGGCAGGGGAACTAC

Cloning into pGEMt-easy vector and sequencing

DNA fragments obtained by PCR using A-tailing Taq polymerase were ligated in the pGEMt-easy vector system [Promega, Fitchburg, USA]. A reaction mixture containing 4.2 μl of gel extracted PCR product, 1 μl pGEMt-easy Vector and 0.8 μl T4 fast ligase within an appropriate buffer were incubated at 16 C for 5-12 hours. Subsequently 1 μl of the ligation mixture was transformed into 100 μl of competent Xl1-blue or DH5α competent *E. coli* cells. The transformed cells were plated onto LB medium containing 80 μl 0.1 M isopropyl thiogalactoside (IPTG) [Sigma-Aldrich, USA] and 40 μl 20 mg/ ml 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (X-gal)[Sigma-Aldrich, USA] which confers blue/ white selection allowing for the detection of integrated insert into the pGEMt-easy vector. White colonies were picked under sterile conditions and grown over-night in 5 ml LB medium with the appropriate antibiotic. Plasmid DNA was extracted using the QIAGEN mini kit following the manufacturer's instructions. Plasmid DNA and the Sp6 and T7 sequencing primers were submitted for sequencing in individual reactions. Sequencing was performed by the Gene-pool genomics facility at the University of Edinburgh.

Sequencing primers:

Name	Sequence [5']
Sp6:	TATTTAGGTGACACTATAG
T7:	TAATACGACTCACTATAGGG

Genotypes of utilised *E. coli* strains

E. coli	
DH5α^{TM}	F-80*lac*ZΔM15 Δ(*lac*ZYA -*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK-, mK+) *pho*A *sup*E44 λ *thi*1 *gyr*A96 *rel*A1
XL1-blue^{TM}	*end*A1 *gyr*A96(*nal*R) *thi*-1 *rec*A1 *rel*A1 *lac gln*V44 F (Tn10 *pro*AB+ *lacIq* Δ(*lac*Z)M15) *hsd*R17(rK- mK+)
DB3.1^{TM}	F- *gyr*A462 *end*A1 *gln*V44 Δ(sr1-*rec*A) *mcr*B *mrr hsd*S20(rB-, mB-) *ara*14 *gal*K2 *lac*Y1 *pro*A2 *rps*L20(Smr) *xyl*5 Δ *leu mtl*1
DH10B^{TM}	FÐ *mcr*A Δ (*mrr-hsd*RMS-*mcr*BC) 80*lac*Z ΔM15 Δ *lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L *nup*G λ-

Preparation and transformation of *E. coli* competent cells

A fresh colony forming unit (cfu) of an *E. coli* strain was used to inoculate a 5 ml LB medium culture grown at 37°C overnight in the presence of the appropriate antibiotic. This was used to inoculate a 500 ml LB medium culture which was grown at 37°C to an OD_{600} of 0.4-0.5. The cells were pre-chilled by swirling the culture flask for 1-2 min in an ice bath and subsequently centrifuged at 2236 g for 10 min at 4°C. The cells were resuspended in 50 ml of ice cold TSS buffer [LB-medium supplemented with 10 % (v/v) PEG 3350, 5% (v/v) DMSO, 1 M $MgSO_4$, pH 6.5] and snap frozen in liquid nitrogen in appropriate aliquots. The cells were stored at -80°C until used for transformation.

For transformation of plasmid DNA a 100 μl competent cell aliquot was thawed on ice. The DNA was mixed in a second tube with 20 μl of 5X KCM buffer [0.5 M KCl, 0.25 M $MgCl_2$, 0.15 M $CaCl_2$] and diluted to a final volume

of 100 μl with ddH_2O. The cells were added to the KCM-DNA solution and incubated for 20 min on ice. A heat shock of the cells was performed by incubation for 5 min at 37°C followed by immediate incubation on ice for a further 2 min. 2 ml of pre-warmed LB-medium was subsequently added to the cells which were then incubated for 1 hour at 37°C. An aliquot of 100 μl was spread onto selective medium containing the appropriate antibiotic within a petri-dish, which was incubated up-side down overnight at 37°C [Walhout et al., 2000].

Preparation and transformation of *A. tumefaciens* competent cells

A fresh colony of an *A. tumefaciens* strain was used to inoculate a 5 ml LB medium culture grown at 28°C overnight in the presence of the appropriate antibiotic. 2 ml of the overnight culture were used to inoculate a 50 ml LB medium culture and grown at 28°C to a OD_{600} of 0.5. The cells were pre-chilled by swirling the culture flask for 1-2 min in an ice bath and subsequently centrifuged at 3,000 g for 10 min at 4°C. The cells were re-suspended in 1 ml of ice-cold $CaCl_2$ [20 mM], dispensed in 100 μl aliquots and frozen in liquid nitrogen. The cells were stored at -80°C until used for transformation.

The competent cells were thawed on ice for 1 hour and 1 μg of plasmid DNA was added to the cells after which the cells were incubated on ice for a further 30 min. The cells were snap frozen in liquid nitrogen and subsequently incubated at 37°C for 5 min. 1 ml of LB medium was added to the cells which were then incubated for 2-4 hours at 28°C. An aliquot of 100 μl was spread onto selective medium containing the appropriate antibiotic and was

incubated up-side down for two to three days at 28°C [Manda YU, personal communication].

2.5 PCR-based methods

All PCR-based methods were performed on a Peltier Thermal Cycler PTC-200 DNA-Engine [MJ Research]. PCR reactions were performed in a 25 μl reaction volume following a general protocol utilising DNA *Taq* polymerase purified in our lab [1μl] by Dr. Yiqin Wang [Pluthero, 1993] or commercial [0.2 μl] DNA *Taq* polymerase, the appropriate buffer containing 7.5 mM $MgCl_2$, 1 μl dNTPs [10 mM; 2.5 mM of each nucleotide] and appropriate primers.

Unless otherwise stated the following PCR conditions were used for the amplification of DNA. Used annealing temperatures were 4°C below the Tm-value provided and for every 1kb expected amplification size 1 min extension time was allowed.

step 1:	94°C	1:00 min
step 2:	94°C	0:40 min
step 3:	appropriate annealing temp.	0:40 min
step 4:	72°C	appropriate extension time: 0:30 - 2:00 min
step 6:	go to step 2: 29 x	
step 4:	72°C	10:00 min
step 5:	4°C	for ever

Genome walking

The genome walking approach was adapted from the Clonetech [USA] Genome-Walker^{TM} methodology. Genomic DNA from *T. cuspidata* [2.5 μg] was digested in individual reactions with 8 μl [20 U/ μl] of the blunt end restriction enzymes *Pvu*II [Roche, Germany], *EcoR*V, *Stu*I, *Nae*I and *Sma*I [New England Biolabs, UK] in a 100 μl volume for 5 hours at 37°C. To each reaction an equal volume of phenol was added followed by centrifugation at 13,000 rpm for 2 min. The last step was repeated using the aqueous layer and an equal volume of chloroform. The DNA was subsequently precipitated by adding two volumes of ice cold 95% EtOH and 9.5 μl 3 M sodium acetate. After vortexing the reaction was centrifuged at 14,000 rpm for 30 min at 4°C, the supernatant was discarded and the pellet was allowed to air dry. This was subsequently dissolved in 20 μl TE buffer [10 mM Tris-HCl pH 8.0 with, 1 mM EDTA]. From each digest 4 μl were used in a ligation reaction to 2 μl of adaptor [Fig.: 2.3] using 0.5 μl of T4 DNA ligase [20 U/ μl] [New England Biolabs, UK] in an 8 μl reaction volume containing the appropriate buffer at 16°C overnight. The reaction was heat inactivated by incubation at 70°C for 5 min and diluted 10/1 in TE buffer. Of each library 1 μl was used in the primary PCR reaction using Ex*Taq* polymerase [Takara Bio. Inc, Japan], and a 1 in 50 dilution of the latter product was utilised in the secondary PCR reaction as a template. For the nested PCR a 1 in 50 diluted product from secondary PCR was utilised as a template. The amplified fragments were subject to agarose gel electrophoresis and after excision and gel purification cloned into the pGEMt-easy vector system.

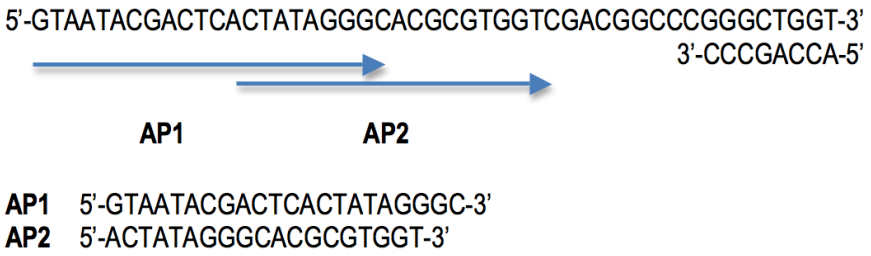

Figure 2.3: The genome walking adapter sequence and the binding sites and sequences of the adapter specific primers AP1 and AP2 [GenomeWalker™; Clonetech, USA].

Name	Sequence [5']
TSgwpri:	GAGGGATATCATCCACAATGGTACTG
TSgwsec:	CACAATGGTACTGGAAGTCTCGGAAAC
TSgwnes:	CTGGAAGTCTCGGAAACCACC

Primary PCR conditions

step 1:	94°C	0:25 min
step 2:	72°C	3:00 min
step 3:	go to step 1 7x	
step 4:	94°C	0:25 min
step 5:	67°C	3:00 min
step 6:	go to step 4 32x	
step 4:	67°C	7:00 min
step 5:	4°C	for ever

Secondary PCR conditions

step 1:	94°C	0:25 min
step 2:	72°C	3:00 min
step 3:	go to step 1 5x	
step 4:	94°C	0:25 min
step 5:	67°C	3:00 min
step 6:	go to step 4 20x	
step 4:	67°C	7:00 min
step 5:	4°C	for ever

Reverse transcriptase-PCR

RT-PCR was perfomed using 1 μg of total RNA from *T. cuspidata*. RNA was extracted using the RNeasy plant RNA extraction kit [Qiagen, Venlo, Netherlands] following the manufacturer's instructions. First strand cDNA synthesis was done utilising the Omniscript® RT kit [Qiagen, Venlo, Netherlands] following the manufacturer's instructions. Reverse transcription was done utilising oligo-d(T) in the presence of the appropriate buffer, 1 μl [5 μM] dNTP and RNasin® RNase inhibitor [40 u/ μl] [Promega, Fitchburg, USA]. The reaction was incubated at 37 °C for 1 hour followed by 10 min at 70 °C. The cDNA was diluted 1 in 10 and utilised as the template for subsequent PCR reactions.

RT-PCR conditions

step 1:	94°C	2:00 min
step 2:	94°C	0:40 min
step 3:	appropriate annealing temp.	0:40 min
step 4:	72°C	0:30 min
step 5:	go to step 2: 20x	
step 4:	72°C	10:00 min
step 5:	4°C	for ever

Primers used for RT-PCR

Name	Sequence [5']
oligo-d(T):	TTTTTTTTTTTTTTTTT
rRNAfwd:	GGGTGTCCCCGCCCCGGAG
rRNAre:	GGGCGTGTTCGGCACGTCCG

Gene specific RT-PCR primers

Name	Sequence [5']
TASY fwd:	AAAAGGCTCGAGGACAACAAG
TASY re:	TTGAATTGGATCAATATAAACTTTC
BAPT fwd:	TTTCTTTGCGTTCTTCCATGATGCG
BAPT re:	GGGAGCGAATGTGTATGGTAGTGCA
PAM fwd:	GGCAGACAACAACGACGCCCT
PAM re:	ACCTACAGTCGCTTCTGCGGAATTTC
DBTNBT fwd:	CTATGGTAATGCCGCTGGTAAT
DBTNBT re:	TTTGTTGAATTACCCCATGTTG

Region of interest (ROI) analysis of RT-PCR

To aid visualisation of the RT-PCR a ROI analysis was performed using a Kodak Digital Science^{TM} 1D setup, which measures the intensity of the amplified RT-PCR band versus the background. The intensity levels were subsequently normalised to the expression level present in CMC at 0 hours post jasmonate treatment (100%).

Cloning of *TED2* probe

Degenerate PCR was used to clone the 412 bp long *TED2* fragment utilised as the probe in northern analysis. The method allows for the amplification of DNA where no sequence information is available. Primer design is performed on the basis of multiple amino acid sequence alignment from homologues of the gene of interest within the same or different organisms. Regions which show a high degree of conservation are chosen for primer design which take sequence variations into account and represent a pool of very similar yet different primer combinations. Primers were designed to amplify the N-terminal region of the *TED2* gene using the aligned authenitic *Zinnia elegans* and two homologues from *Populus trichocarpa* as a sequence guide [Appendix I]. PCR was performed using *T. cuspidata* cDNA as a template. The amplified PCR fragment was gel extracted after agarose electrophoresis and cloned into pGEMt-easy vector for sequence verification. The nucleotide sequence of the isolated *T. cuspidata TED2* fragment was translated into an amino acid sequence which showed 69% sequence identity and 93% sequence similarity to the authentic *TED2* gene from *Z. elegans*.

***TED2* degenerate PCR primers**	
forward:	ACNGTVTCNACNAARGARAARGC
reverse	CNGCYTGNGANARNGGRTA
	For mixed base defentition see Appendix I

***TED2* PCR conditions**

step 1:	94°C	0:10 min
step 2:	94°C	0:20 min
step 3:	57°C	0:20 min
step 4:	72°C	0:30 min
step 5:	go to step 2: 30x	
step 4:	72°C	10:00 min
step 5:	4°C	for ever

2.6 Northern blot analysis

Total RNA was extracted from *T. cuspidata* csc using the Quiagen plant RNase extraction kit [Qiagen, Venlo, Netherlands] following the manufacturer's instructions. The absorbance of each sample was measured at 260 nm, and used to calculate the concentration of RNA. Samples [10 μg] were separated on formaldehyde-agarose gels [Sambrook et al., 1989], transferred to a Hybond TM-N hybridisation membrane according to the instructions of the supplier [Amersham, UK] and hybridized with the relevant probe. Blots were washed twice for 30 min each at 65°C in 4 x SSC, 1% (w/v) SDS, which was followed by two washes at 65°C in 4 x SSC, 0.5% (w/v) SDS. Blots were exposed to X-Omat-ARTM imaging film (Kodak) for an appropriate period. The previously cloned *TED2* probe was excised using the *Eco*RI restriction enzyme and labelled with α-^{32}P-dCTP by random priming using the Prime-a-Gene labelling system [Promega, UK].

2.7 Genomic library construction

Isolation of nuclei from *T. cuspidata*

Isolation of nuclear plugs was performed as previously described [Hein et al., 2005]. 20 g of frozen *T. cuspidata* csc were ground to a fine powder with liquid nitrogen and transferred to a flask containing 350 ml MEB buffer, pH 6.0 [1 M MPD (2-methyl-2,4-pentanediol); 10 mM PIPES KOH; 10 mM $MgCl_2$; 4% (w/v) polyvinylpyrrolidone-10; 10 mM sodium metabisulfite; 0.2% (v/v) β-mercaptoethanol; 0.5% (v/v) Triton X-100; the pH was adjusted with HCl]. The homogenate was incubated at room temperature for 12 min with a gentle stir every 2 min and was subsequently filtered through one layer of miracloth and then again through two layers of miracloth. Small quantities of the homogenate were filtered by gravity and gentle agitation through a stack of 40 μm nylon mesh on top of a 20 μm nylon mesh. The homogenate was centrifuged at 650 g for 20 min and the pelleted nuclei were resuspended in 15 ml of MPDB buffer, pH 7.0 [MEB buffer without the polyvinylpyrrolidone-10; the pH was adjusted with NaOH]. The homogenate containing the nuclei was layered on a discontinuous percoll gradient consisting of 3 ml layers within a glass tube of 95%, 60%, 45%, 30% and 15% percoll diluted with MPDB. Following centrifugation at 650 g for 20 min the nuclei were harvested from the interphase below the 30% percoll layer, pooled and diluted with 15 ml MPDB buffer. The nuclei suspension was subsequently loaded on top of 10 ml 85% percoll diluted with MPDB, centrifuged at 650 g for 20 min and harvested from the percoll/ MPDB buffer interphase and diluted in 20 ml of MPDB. The last

step was repeated with 37.5% percoll. The supernatant was gently removed and the collected nuclei at the bottom of the tube were resuspended in 25 ml MPDB buffer. This step was repeated twice again where in the last repeat the nuclei were resuspended in MPDB(-) buffer not containing β-mercaptoethanol and Triton-X [Hein et al., 2005].

Embedding of nuclei, lysis and washing of nuclear plugs

Embedding of nuclei was performed as previously described [Peterson et al., 2002]. InCert® agarose [Cambrex, USA] was utilised for the preparation of nuclear plugs. A 1 % agarose solution was prepared in MPDB(-) buffer and incubated at 45°C in a waterbath. The nuclei were pre-warmed to 45°C for 5-10 min and gently mixed with an approximately equal volume of the agarose solution using a wide-bore pipette tip (with cut off end). The mixture was carefully aliquoted into pre-chilled disposable plug moulds [CHEF, Bio-Rad, USA].

Lysis and washing of nuclear plugs was performed as previously described [Farrar and Donnison, 2007]. The nuclear plugs were allowed to set overnight at 4°C. For subsequent lysis the plugs were gently retrieved from the mould and incubated in 50 ml lysis buffer [0.5 M EDTA pH 9, 1 % (w/v) sodium lauryl sarcosine, 0.1 mg/ ml proteinase K] at 50°C for 24-48 hours with gentle shaking. Following this the plugs were washed in 0.5 M EDTA [pH 9] for 1 hour at 50°C with gentle shaking. The last step was repeated twice utilising

0.05 M EDTA [pH 8]. Following this the plugs were washed twice more for 1 hour in TE (20:50) buffer [20 mM Tris-HCl, 50 mM EDTA, pH 8]. Prior to enzymatic digestion the nuclear plugs were washed three times for 1 hour each at 4°C in ice cold TE-PMSF buffer [10 mM Tris-HCl, 1 mM EDTA pH 8, 0.1 mM PMSF] which was repeated a further three times utilising TE (20:50) buffer.

Electroelution of HMW DNA

Electroelution of HMW BAC insert DNA was performed as previously described [Strong et al., 1997]. The dialysis tubing (12,000 - 14,000 D. molecular weight cut off) [Medicell Int. Ltd., UK] was prepared by heating the strips [8-10 cm length] at 90°C for 10 min in buffer A [1 mM EDTA 2 % (w/v) Sodium bicarbonate] followed by boiling the membrane for 10 min in dH_2O. The tubing was subsequently rinsed in dH_2O several times and stored at 4°C in 50 % EtOH. Immediately prior to use the dialysis tubing was rinsed thoroughly in sterile dH_2O and then in sterile 1 X TAE buffer [40 mM Tris-acetate, 1 mM EDTA, pH 8]. For each electroelution, a gel slice of 300 mg was used and prepared by equilibration in 50 ml of 1 X TAE buffer at 4°C for 30 min. The gel slice was placed longitudinally into the membrane and 300 - 400 μl of sterile 1 X TAE was added. Care was taken that upon sealing of the dialysis tubing all air was removed. The dialysis tubing containing the membrane was submerged in 4°C cold 1 X TAE buffer in an electrophoration tank with the membrane being proximal to the negative pole. Electroelution was performed at 4°C at 4-5 V/ cm for 2 hours. To disassociate the DNA from the mem-

brane the polarity was reversed for one minute. The eluted DNA was carefully pipetted from the dialysis tubing.

BAC prep alkaline lysis

DNA was isolated from BAC clones as as previously described [Farrar and Donnison, 2007]. 10 ml of LB medium containing the appropriate antibiotic was inoculated with a fresh cfu and grown overnight at 37°C. The culture was sequentially centrifuged for 5 min at maximum speed in a 2 ml eppendorf tube and the supernatant was discarded. The pellet was subsequently resuspended in 200 μl of ice cold buffer I [25 mM Tris-HCl, pH 8.0; 10 mM EDTA; 0.9 % glucose] and chilled on ice. 200 μl of freshly prepared buffer II [1% SDS; 0.2 M NaOH] were added to each sample, mixed thoroughly by inversion and subsequently incubated on ice for 4 min. 300 μl of solution III [3 M KOAc; 11.5% glacial acetic acid] were added and mixed thoroughly. After incubation for 1 hour on ice, the samples were centrifuged at maximum speed for 10 min and the supernatant was discarded. An equal volume of chilled isopropanol was added to each sample and mixed by inversion. The samples were subsequently centrifuged at maximum speed for 10 min and the supernatant was discarded. The pellet was washed with 70% EtOH, air dried and resuspended in 50 μl dH_20.

For the determination of BAC insert size, 20 μl of the BAC DNA prep from 19 BAC clones were digested with 0.3 μl of the restriction enzyme *Not*I at 37°C for 3 hours. The BAC DNA restriction reactions were subject to PFGE and the average insert size was determined according to the employed

size marker.

2.8 Web resources

Function	Web-adress
Plant *Cis* element database:	http://www.dna.affrc.go.jp/PLACE/
Plant TF database:	http://plntfdb.bio.uni-potsdam.de/v3.0/
NCBI blast:	http://blast.ncbi.nlm.nih.gov/Blast.cgi
NCBI ORF finder:	http://www.ncbi.nlm.nih.gov/projects/gorf/
NCBI conserved domain search:	http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
Base translation:	http://expasy.org/tools/dna.html
Sequence alignment:	www.ch.embnet.org/software/TCoffee.html
TAIR N-browse:	http://www.arabidopsis.org/tools/nbrowse.jsp

Chapter 3

Cloning and analysis of Taxol^{TM} biosynthetic gene promoters

3.1 Introduction

Biosynthesis and accumulation of secondary metabolites, such as the taxene diterpenoid Taxol^{TM}, is induced by biotic- and abiotic-elicitors [Zhang et al., 2000]. This results in the expression of Taxol^{TM} biosynthetic enzymes which are mediated by the plant hormone methyljasmonate and controlled at the transcriptional level [Nims et al., 2006]. Transcription is governed by a complex network of regulators which control promoter activity in response to signal transduction cascades induced by environmental cues [Singh, 1998]. Several TFs are often required to interact with *cis* elements in the promoter region of target genes to activate transcription [Endt et al., 2002].

Despite the growing demand for the anticancer drug Taxol^{TM} relatively

little is known about the transcriptional regulation of the biosynthetic genes involved in the Taxol™ pathway. The latter is partly due to the fact that the Taxol™ pathway is not yet fully characterised [Nims et al., 2006]. Furthermore, as a non-model organism the lack of genomic sequence data for *Taxus* species poses another limitation on work concerned with DNA out-with the transcribed population. The genome of *Taxus cuspidata* is very large with 10,850 Mb. To restrict the target size, functional studies used exclusively cDNA cloned from RNA of *Taxus* [Guo et al., 2006].

Promoter studies in this thesis concentrate on the genes encoding TASY, BAPT and DBTNBT. The terpene synthase TASY is the committing enzyme in the Taxol™ pathway and as such, synthesising the taxene core from the precursor molecule IPP [Wildung and Croteau, 1996], might be governed by transcriptional regulation which tightly influences the expression of downstream pathway genes.

The biosynthesis pathway of diterpenoids does not conform to a linear fashion but with pathway branches feeding in or branching off at multiple steps resulting in different intermediates and end products, which is also true for the Taxol™ pathway [Nims et al., 2006] [reviewed in 1.5.2; Fig.: 1.4]. BAPT and DBTNBT are enzymes responsible for the last steps in Taxol™ biosynthesis. BAPT catalyses the attachment of the side chain precursor to the taxene core [Walker et al., 2002] and the last benzoylation is catalysed by DBTNBT [Walker et al., 2002].

Hence the latter enzymes may represent biosynthetic steps unique to the Taxol™ pathway which by extension may also apply for the underlying

transcriptional regulation. Furthermore DBTNBT was identified as a possible rate limiting step within the pathway exhibiting a low expression level [Nims et al., 2006].

Cloning and analysis of Taxol^{TM} biosynthetic gene promoter regions will greatly contribute to the identification of transcriptional regulators and the underlying molecular mechanism involved.

3.2 Genomic library construction attempts

In order to clone the promoter regions of *TASY*, *BAPT* and *DBTNBT* attempts have been made to construct and screen a genomic DNA library from *T. cuspidata*. Plant BAC libraries have been constructed from a wide variety of species [Peterson et al., 2002] and are valuable resources providing tools for functional gene elucidation within the genomic context [Farrar and Donnison, 2007]. Genes encoding for enzymes involved in plant secondary metabolism have in several instances been shown to be linked. Biosynthetic genes of diterpene biosynthesis in rice and triterpene biosynthesis in oat and *A. thaliana* are located in gene clusters [Osbourn, 2010]. This finding provided a further rational for the construction of a BAC library which potentially aided to the discovery of missing Taxol^{TM} pathway components.

For the isolation of Megabase-size genomic DNA from *T. cuspidata* nuclear plugs were prepared as previously described by Hein et al., [2005] and embedded in 0.8% low melting point agarose [Sigma-Aldrich Co., Missouri, USA]. The agarose provides a protective matrix preventing the damage of nuclei and

thus shearing of the genomic DNA. Furthermore the porous characteristic of the agarose allows for the diffusion of reagents necessary for purification and manipulation of the genomic DNA in subsequent steps [Schwartz and Cantor, 1984]. To make the genomic DNA accessible for enzymatic restriction the nuclear plugs were lysed and washed as described by Farrar and Donnison [2007]. Prior to partial digestion with the type II restriction endonuclease *Hind*III [New England Biolabs, UK] a test restriction was carried out to determine the optimal enzyme concentration. For this a quarter of a nuclear plug was incu-

BAC *Hind*III test restriction of *T. cuspidata* HMW genomic DNA

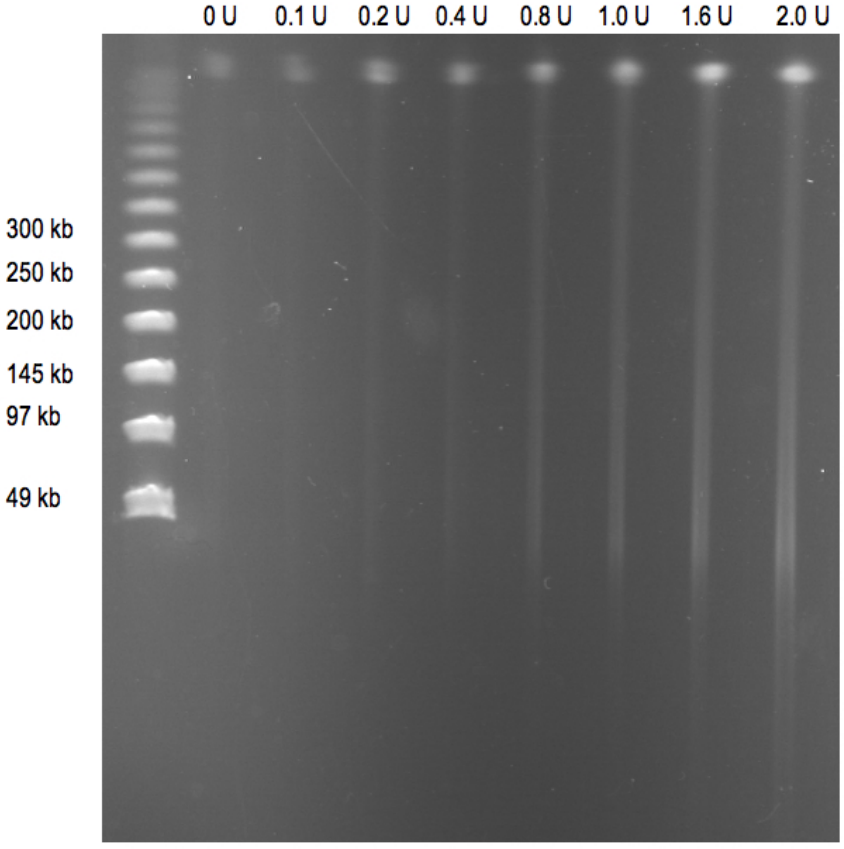

Figure 3.1: One quarter nuclear plug was partially digested for 1 h with increasing units (U) of *Hind*III and analysed by PFG-electrophoresis.

bated for four hours at 4°C with the enzyme mix to enable complete diffusion of the reagents necessary for restriction digestion of the genomic DNA. Included in the individual enzyme mixtures was the buffer and various concentrations of the restriction enzyme *Hind*III which correlated to 0, 0.1, 0.2, 0.4, 0.8, 1.0, 1.6 and 2.0 enzyme units [Fig.: 3.1]. The enzyme concentration of 1.0 unit was found to be optimal for partial digestion generating a substantial portion of genomic DNA of the target size of 100-300 kb and at the same time showing a reduced amount of undigested DNA thus resulting in a larger overall volume of

First and second size selection of *T. cuspidata* HMW genomic DNA

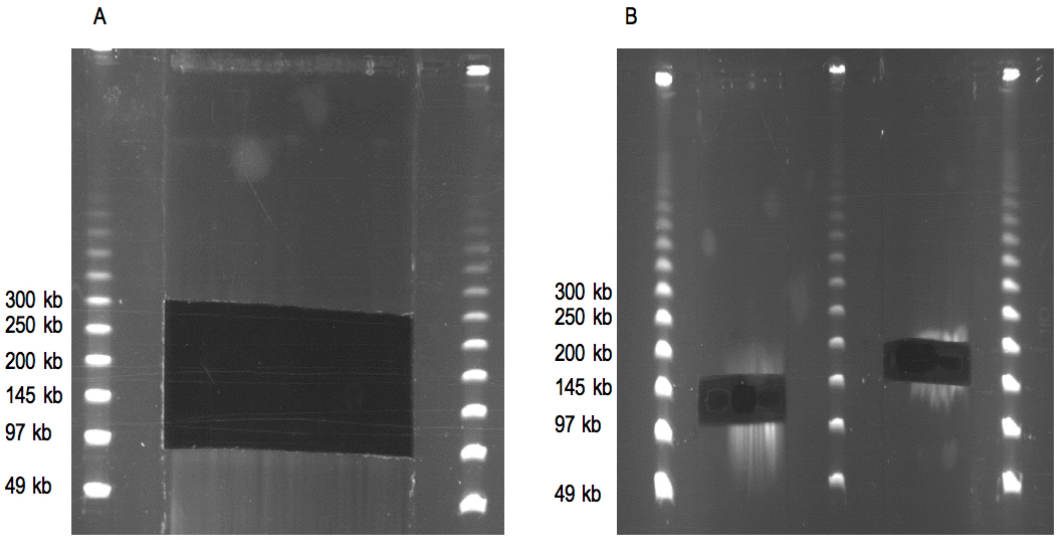

Figure 3.2: Re-assembled gel fragments after PFG-electrophoresis and ethidium bromide staining; [A]: First size selection of the target size of 100-300 kb. [B]: Two gel segments containing HMW DNA from 100-200 kb and 200-300 kb were subject to a second size selection. Visible above and below the excised gel fraction is the co-migrating HMW DNA from the first size selection out-with the target size.

partially digested HMW genomic DNA. The nuclear plugs were subsequently incubated with the enzyme mix as described above and individually digested with the previously determined *Hind*III concentration of 1.0 unit at 37°C for 1 hour. The restriction enzyme *Hind*III was chosen because of the employed BAC vector, pIndigoBAC-5 [Epicentre Biotechnologies, Madison USA] carries compatible *Hind*III cloning sites. After digestion the nuclear plugs were subject to PFG electrophoresis using a CHEF-DRII setup [Bio-Rad Laboratories Inc., California USA] on a 1% agarose gel [Seakam gold, Cambrex Bio Science, Maine USA].

In order to isolate the HMW genomic DNA of 100-300 kb, two size selections were performed [Fig.: 3.2]. Prior to excision of target size DNA, the left and right edges, which contained the co-migrating DNA size marker within the agarose gel, were removed and stained with ethidium bromide [Sigma-Aldrich Co., Missouri, USA]. These were subjected to UV-light and physical markers were cut into the agarose gel highlighting the position of the appropriate size marker DNA bands. The agarose gels containing the DNA size marker were reassembled with the previously removed middle part of the gel containing the partially digested genomic DNA and according to the physical markers within the left and right gel fragments the appropriate gel section was excised [Fig.: 3.2 A].

In order to exclude co-migrating smaller genomic DNA fragments outwith the target size a second size selection was performed. This included PFG electrophoresis of the previously excised gel fragments. Digested DNA from 100-200 kb and 200-300 kb were individually purified in the same way as

described above [Fig.: 3.2 B].

Following the second size selection, the agarose gel embedded DNA is liberated by electroelution, which is a electrophoresis coupled dialysis of the DNA, as previously described by Strong et al., [1997]. The concentration of the partially digested *Hind*III genomic DNA was subsequently determined [3.5 ng/μ] [Fig.: 3.3 A] using a nano-drop setup.

Isolated HMW genomic DNA for BAC library construction and the BAC vector

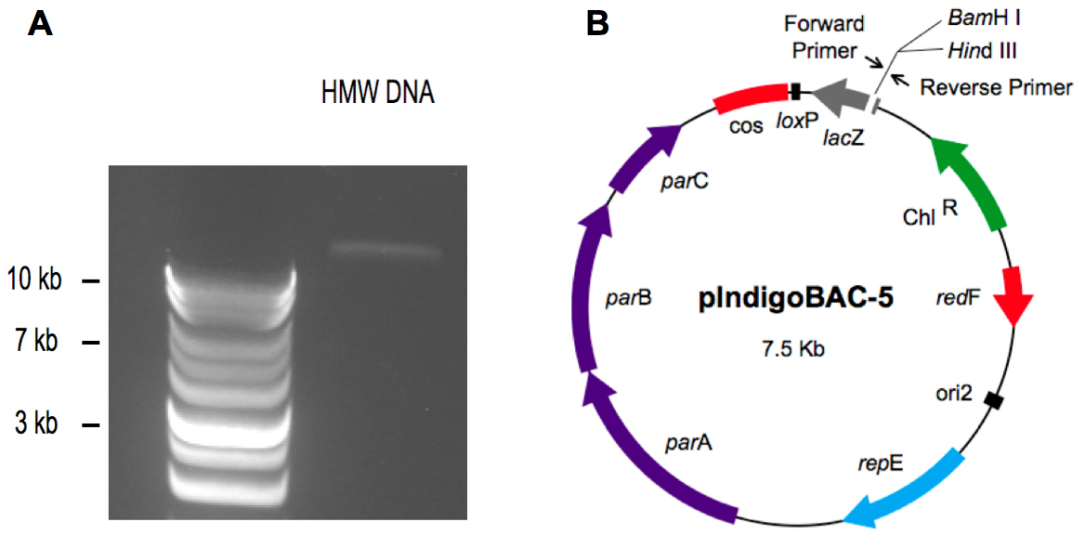

Figure 3.3: [A]: An aliquot of the pooled electroeluted HMW genomic DNA for BAC library construction from both size selected fractions (100-200 and 200-300) was run by standard electrophoresis on a 1% agarose gel; [B]: The pIndigoBAC-5 vector used for construction of the *T. cuspidata* genomic library. The Vector is *Hind*III precut and dephosphorylated thus ensuring high BAC insert integration efficiency.

The subsequent ligation of the HMW DNA was performed as described by Peterson et al., [2002], 140 ng of partially digested genomic DNA and 50

ng of the pre-cut and dephosphorylated BAC vector pIndigoBAC-5 [Fig.: 3.3 B] were incubated using T4 ligase [400.000 units/ml, New England Biolabs, UK] over night at 16°C. The pre-cut vector was utilised to ensure high insert integration efficiency. Following ligation a dialysis was performed.

For transformation of the BAC constructs electrocompetent *E. coli* cells [*E. cloni* 10G BAC cells, Lucigen, Wisconsin USA] were utilised. This *E. coli* strain, *DH10B*, carries the *mcr* and *mrr* mutations thus allowing for the transformation of large methylated genomic DNA fragments without generating deletions or rearrangements of the insert. Furthermore, these cells were optimised for high transformation efficiency.

With a haploid genome size of 10,850 Mb a target of 400,000 cfus are required with an average insert size of 110 kb to reach 4x genome coverage. The competent cells employed exhibited an expected transformation efficiency of $1x10^7$ cfu/μg 150 kb BAC insert [Lucigen, Wisconsin USA] per 20μl cell aliquot, however transformation with 32 ng of the ligation mixture resulted in 3601 cfus which correspond to 1.2% of the expected outcome. In order to increase the transformation efficiency an increased amount of 200 ng of BAC insert DNA was utilised which resulted in 2900 cfus after transformation. Further attempts to increase the ligation efficiency were made by using high strength T4 ligase [2,000,000 units/ml] however, no increase in cfus was observed. In order to circumvent the need to buy electro-competent cells, home-made competent cells were utilised [described in 2.4], however this resulted in a significant decrease of 150-1500 cfus per transformation using a 20 μl cell aliquot with an average of 1110 cfus after several transformations.

In order to determine the average insert size of the cloned BAC constructs, liquid cultures were grown from 19 cfus over night at 37°C and following a BAC prep by alkaline lysis a restriction digest was performed to determine the average insert size. For this the endonuclease *Not*I was chosen as its recognition sequence motif flanked the insertion sites on the BAC vector. Furthermore *Not*I recognises an eight nucleotide sequence motif which occurs less frequently within the genomic DNA thereby generating larger fragments or no restriction within the BAC insert. This is favoured when determining the BAC insert size. Restriction digest with *Not*I showed an average insert

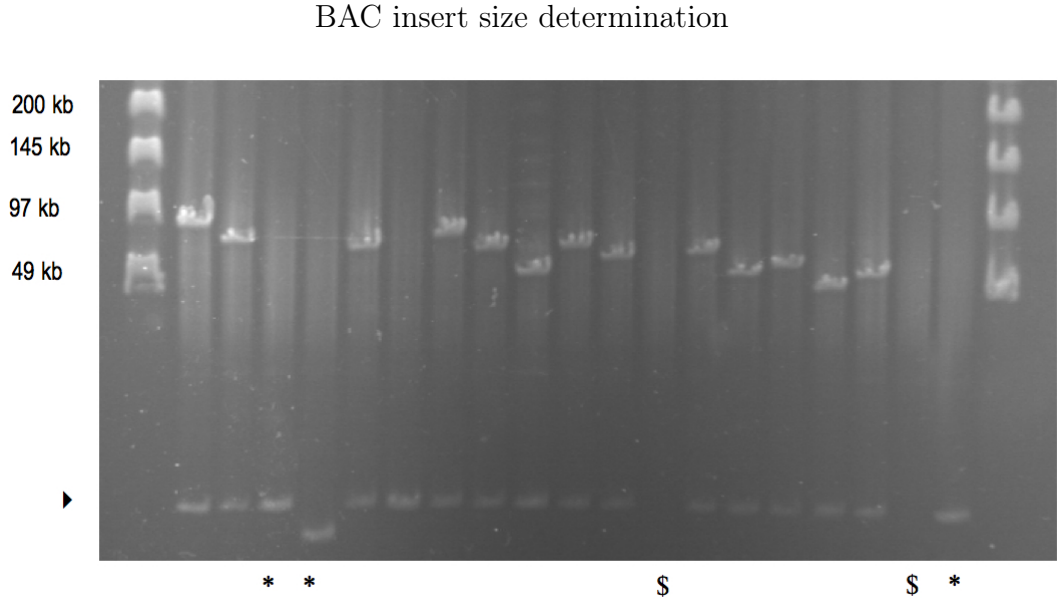

Figure 3.4: DNA from BAC clones was isolated and digested with *Not*I to determine the average insert size. The arrow below the co-migrating size marker indicates the pIndigoBAC-5 vector backbone, asterisks indicates an empty vector and the dollar symbol a colony forming unit (cfu) not harbouring a BAC construct. An average insert size of 73 kb was determined.

size of 73 Kb with 76% of the cfus harbouring a BAC construct [Fig.: 3.4].

Recalculation of the required cfus, the smaller than expected insert size along with the absence of a BAC construct in 24% of cfus, after, resulted in attainment of the target of 4x genome coverage of about 793.000 cfus. Due to economical and logistical constraints, concerned with the amount of required BAC vector, insert DNA, competent cells and storage, the construction of the genomic library from *T. cuspidata* was abandoned.

3.3 Genome walking

The Genome walking approach was utilised in order to clone Taxol^{TM} biosynthetic gene promoters which resulted in the isolation of a 513 bp promoter fragment from *TASY*.

The genome walking approach includes the individual partial digestion of *T. cuspidata* genomic DNA with restriction endonucleases generating a blunt end restriction site. Enzymes used for digestion include *Pvu*II [Roche, Germany], *EcoR*V, *Stu*I, *Nae*I and *Sma*I [New England Biolabs, UK]. As recognition sequences for the different restriction enzymes occur at different locations within the genome a variety of enzymes were used to increase the likelihood of generating a partially digested genome-adapter fragment which would be favoured by PCR and result in amplification of the genomic fragment of interest. A blunt-end adaptor [described in 2.5; Fig.: 2.1] was ligated to the different libraries of partially digested genomic DNA. A primary PCR using ExTaq polymerase [Takara Bio. Inc, Japan] was performed using a gene spe-

cific primer [approximately 200 bp downstream of the ATG] and an adaptor specific primer, AP1, utilising the partially digested genome walking libraries as a template [Fig.: 3.5A]. The amplified DNA was diluted 1:50 in TE buffer and an appropriate volume was used as a template in a secondary PCR [Fig.: 3.6], which utilised another gene specific primer, partly overlapping with the previously used, and a nested adaptor specific primer AP2. Following this the amplified products from the *Pvu*II were again diluted 1:50 in TE buffer which was used in a nested PCR using the adaptor specific primer AP2 and a further nested gene specific primer. Nested PCR was performed to verify the authenticity of the amplified fragment using the *TASY* specific primers.

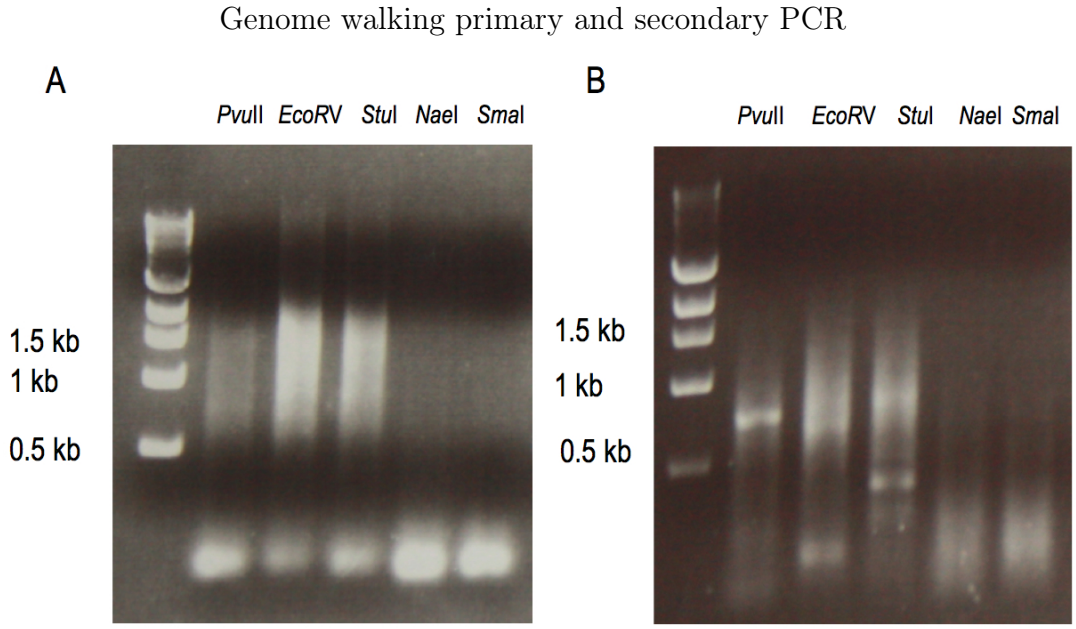

Figure 3.5: A primary [A] and secondary [B] PCR was performed on adapter ligated libraries of *T. cuspidata* partially digested genomic DNA using various blunt end restriction endonucleases.

The amplified fragment from the nested PCR [Fig.: 3.6] was subject to gel-electrophoresis using a 0.8% agarose gel and after excision the DNA was extracted using the gel extraction kit [Qiagen, Venlo, Netherlands] according to the manufacturer's instructions.

The amplified fragment was subsequently cloned into the pGEMt-easy vector system which involves the T4 ligase mediated insertion of the 5'-A-tailed PCR product into the 3'-terminal thiamidine flanked linearised vector. Subsequent sequencing using the T7 and SP6 sequencing primers whose annealing sites are present on the pGEMt-easy vector confirmed the isolation of a 700 bp fragment [Fig.: 3.5] which corresponds to a 513 bp long promoter region of *TASY*.

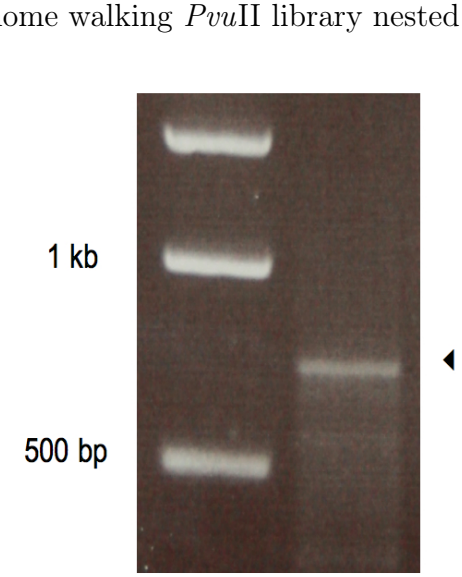

Figure 3.6: A 700 bp fragment (arrow) was amplified in the nested PCR of the genome walking approach which includes 513 bp of the *TASY* promoter region.

Although this procedure was also performed for the Taxol^{TM} biosynthetic genes *BAPT* and *DBTNBT*, the amplified fragments [data not shown] after sequencing could not be confirmed as authentic promoter fragments of the genes of interest.

Recently the promoter sequences from *TASY* (1436 bp) and *DBTNBT* (1891 bp) were submitted to the NCBI nucleotide public domain. Sequence alignment of the 1436 bp long promoter fragment with the sequence obtained through genome walking showed an overlap which further confirmed the authenticity of the cloned *TASY* promoter.

3.4 In-silico promoter analysis

In-silico analysis of the *TASY* and *DBTNBT* promoter sequences was performed. In order to estimate the transcription start site, the promoter sequences were compared with the cDNA entries in the public domain. This identified a 5' UTR of 21 and 35 nucleotides upstream of the translation start site [ATG] for the *TASY* and *DBTNBT* promoters respectively.

The in-silico analysis concentrated on the identification of known jasmonate and pathogen-responsive elements. Also included were target binding sites of TF families previously shown to be involved in transcriptional regulation of secondary metabolites.

Several *cis* elements which may confer jasmonate and pathogen inducibility were located within the proximal *TASY* promoter region [Table: 3.1]. A jasmonate-responsive GCC-box element is located at position -142 [according

to the transcription start site]. A G-box and T/G box at positions -158 and -255 respectively and a CGTCA motif at position -189 are also present in the proximal region.

cis element	Sequence	Position	Direction
GCC-box	GCCGCC	-142	forward
G-box	CACGTG	-158	forward
T/G-box	AACGTG	-255	forward
CGTCA-motif	CGTCA	-189	forward
W-box	TGACT	-136	forward
	TGACT	-352	reverse
	TGACT	-801	forward
	TGACT	-904	forward
	TGACC	-1304	forward
MYB I cons.	TAACTG	-1295	reverse
	TAACGG	- 1329	forward

Table 3.1: *TASY* promoter *cis* elements

Within the proximal promoter region is a pathogen-responsive W-box element at position -139. It is noteworthy that this overlaps partly with the GCC-box element. However four further W-boxes are located within the *TASY* promoter. Furthermore two MYB I consensus motifs are present at positions -1295 and -1329.

In-silico analysis of the *DBTNBT* promoter revealed a G-box motif at position -92 in the proximal promoter region and a G-box related element at position -443 [Table: 3.2]. Also present are two W-boxes one of which is

located within the proximal promoter at position -146. Furthermore several MYB consensus motif binding sites of type I and II are present in the *DBTNBT* promoter region.

cis element	Sequence	Position	Direction
G-box	CACGTG	-92	forward
	CACGTT	-443	reverse
W-box	TGACT	-146	reverse
	TGACC	-923	reverse
MYB I cons.	CCGTTG	-722	forward
	CTGTTG	-1139	forward
	CGGTTG	-1726	reverse
MYB II cons.	TAACAAA	-261	forward

Table 3.2: *DBTNBT* promoter *cis* elements

3.5 Functional characterisation of the *TASY* and *DBTNBT* promoters

Particle bombardment of *Taxus cuspidata*

To assess the jasmonate-responsiveness of the cloned *TASY* promoter fragment attempts have been made to develop a transient transformation assay using the utilised *T. cuspidata* cell suspension culture [csc]. Initial studies concentrated on particle bombardment of the utilised plant material which involves the shooting of microcarriers [gold particles] which deliver the desired DNA into the plant cell [Gan, 1989]. Exogenous DNA once it enters the nucleus can be

transcribed and expressed by the host plant [Levy et al., 2005].This extensively used approach for transient transformation of a variety of different plant species provides a platform for the expression of chimeric gene constructs and their analysis [Klein et al., 1992].

Transiently transformed *T. cuspidata* cells expressing the reporter *GFP*

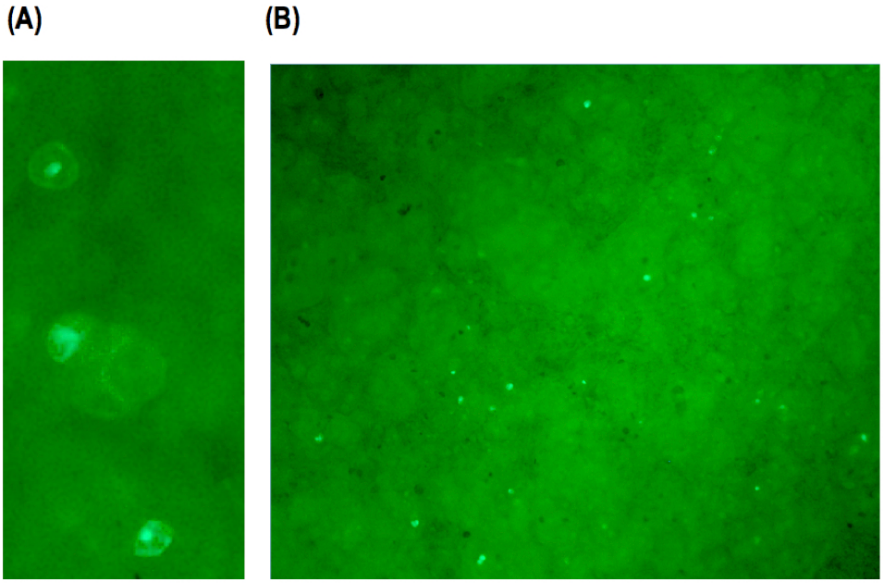

Figure 3.7: *T. cuspidata* GFP activity. [A]: Three individual cells showing GFP fluorescence under blue light 14 hours after bombardment; [B]: Overview of cells expressing the GFP reporter.

Bombardment of the *T. cuspidata* csc was performed with the Bio-Rad PDS1000/He system [Bio-Rad, Hercules, CA] according to the manufacturer's instructions. In order to asses the efficiency of the transient transformation conditions a plasmid harbouring the *GFP* reporter gene controlled by the constitutive *CaM V 35S* promoter was utilised. A thin layer of *T. cuspidata* csc was transferred onto a sterile filter paper four hours prior to bombardment and placed onto solid growth medium containing 8 g/L agar-agar. For each

shot 0.5 μg of plasmid DNA was precipitated onto 500 μg of 1.6 μm gold microcarriers.

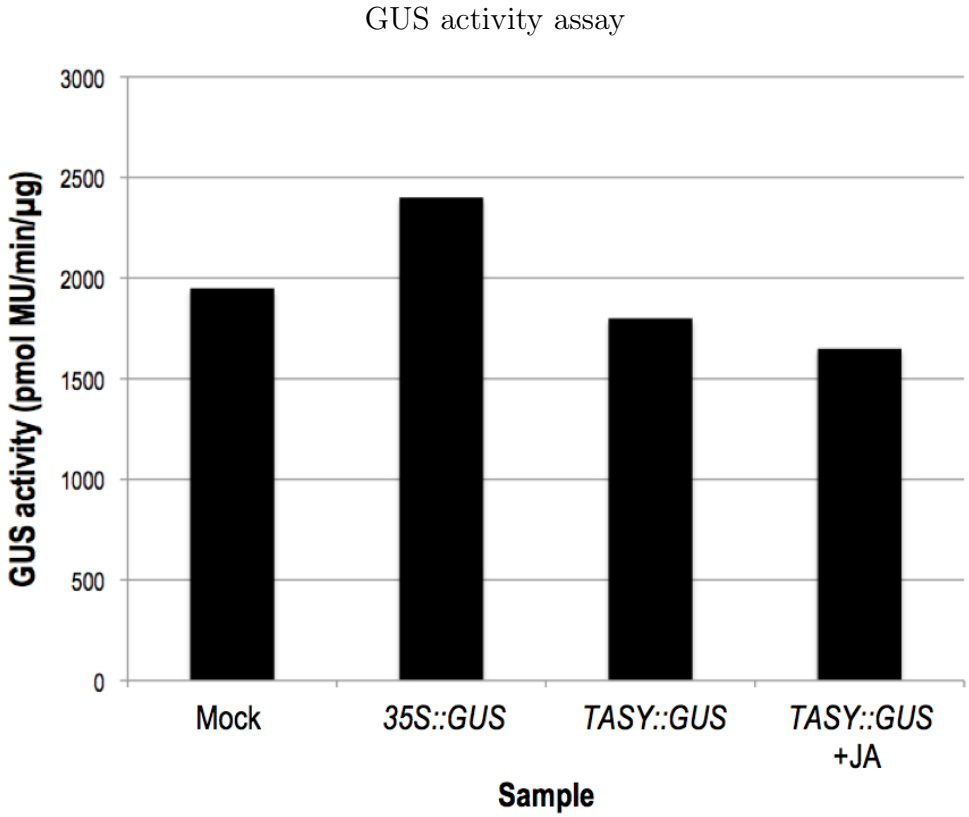

Figure 3.8: *T. cuspidata* cells were transfromed by particle bombardment with a construct harbouring the -513 *TASY* promoter construct driving the reporter gene *GUS* with and without methyljasmonate. Mock: promoterless *GUS* plasmid

Rupture disc pressures of 650 psi, 1100 psi and 1350 psi were applied in combination with 1.0 μm and 1.6 μm microcarriers. No *GFP* expression was observed with 650 psi and 1100 psi using 1.0 μm microcarriers and only limited *GFP* expression with 1350 psi. The latter rapture disc pressure was therefore used with an increased size of 1.6 μm microcarriers which increased the number of *GFP* expressing *T. cuspidata* cells. Effects of various target

distances of 9 cm, 6 cm and 3 cm were tested with the latter showing the highest *GFP* expression. In an effort to further increase the transformation efficiency, three replicate shots were applied which resulted in a larger number of transformed cells [Fig.: 3.7].

A promoter *GUS* fusion construct was cloned in order to asses the methyl-jasmonate inducibility of the -513 *TASY* promoter. Bombardment was performed as described above with three sequential shots using 1.6 μm microcarriers and a rapture disc pressure of 1350 psi. After bombardment the cells were harvested from the filter paper and transferred into liquid growth medium. A population of the bombarded cells was elicited with 100 μM MeJA. The cells were harvested after 24 hours which correlates to the peak of *TASY* expression after elicitation [Nims et al., 2006] and a GUS activity assay was performed. However, no expression of *GUS* was detected [Fig.: 3.8].

Taxus cuspidata protoplast isolation and transformation

Due to the limited success of particle bombardment, attempts have been made to utilise electroporation of isolated protoplasts from *T. cuspidata* as a transient transformation platform. Protoplasts are plant cells which had their cell walls removed, however, they retain their viability, cell identity and differentiated state [Sheen, 2001]. Transformation of protoplasts has been shown in various plant species where they have been utilised to study a variety of different cellular aspects [Davey et al., 2005, Sheen, 2001].

To determine the ideal enzyme composition for the digestion of the cell wall present in *T. cuspidata* procambium csc, several enzymes were utilised.

Cellulases are widely used in the isolation of protoplasts [Cocking, 1972]. Three different commercial cellulase preparations from *Trichoderma viride* were utilised, cellulase-*Trichoderma* [Calbiochem, USA], cellulase Onozuka RS and cellulase Onozuka R10. Equal settled cell volumes of 1 ml of the *T. cuspidata* csc were digested with 1% [w/v] cellulase (1U/mg) and equal concentrations of Onozuka RS and cellulase Onozuka R10 dissolved in TEX buffer over night in the dark at room temperature. Cellulase-*Trichoderma* was found to be most effective. Further enzymes used in the isolation of *T. cuspidata* protoplasts were 0.1% [w/v] pectolyase Y-23 [MP Biochemicals, UK] and 0.2% [w/v] Driselase^{TM} [Sigma-Aldrich, USA]. Protoplast isolation was performed in an osmoticum of 0.5 M Mannitol [Naill and Roberts, 2005b].

To determine electroporation conditions suitable for protoplasts isolated from the *T. cuspidata* procambium csc a voltage range from 140 V to 200 V in 15 V increments was employed. 500 μl of previously isolated protoplasts with an estimated density of 3 X 10^6 [according to the thickness of the isolated protoplast layer] were mixed with 50 μg of plasmid DNA harbouring the α-*AMYLASE* reporter gene under the control of the *CaMV* 35S promoter. Following electroporation the α-amylase reporter assay was performed. The highest transformation of *T. cuspidata* procambium protoplasts was achieved at a 170 volts [Fig.: 3.9]. The expected OD of the reporter assay is between 0.1 and 1.0 with lower readings being not reliable [Prof. J. Denecke, personal communication]. Although an isolation procedure for protoplasts from *T. cuspidata* procambium cells has been established the observed transformation efficiency by protoplast electroporation was insufficient.

Agrobacterium infiltration of *Nicotiana benthamiana*

Although attempts have been made to utilise the plant model *T. cuspidata* for transformation and analysis of the *TASY* and *DBTNBT* promoters, an efficient transient transformation approach could not be established. Thus the chimeric constructs harbouring the *TASY* and *DBTNBT* gene promoters [Fig.: 3.10 A and B] driving the *GUS* reporter gene were transformed into *Agrobacterium tumefaciens* and used for infiltration of *Nicotiana benthamiana* leaves.

The constructs were cloned using the Gateway® vector pMDC162 [Fig.:

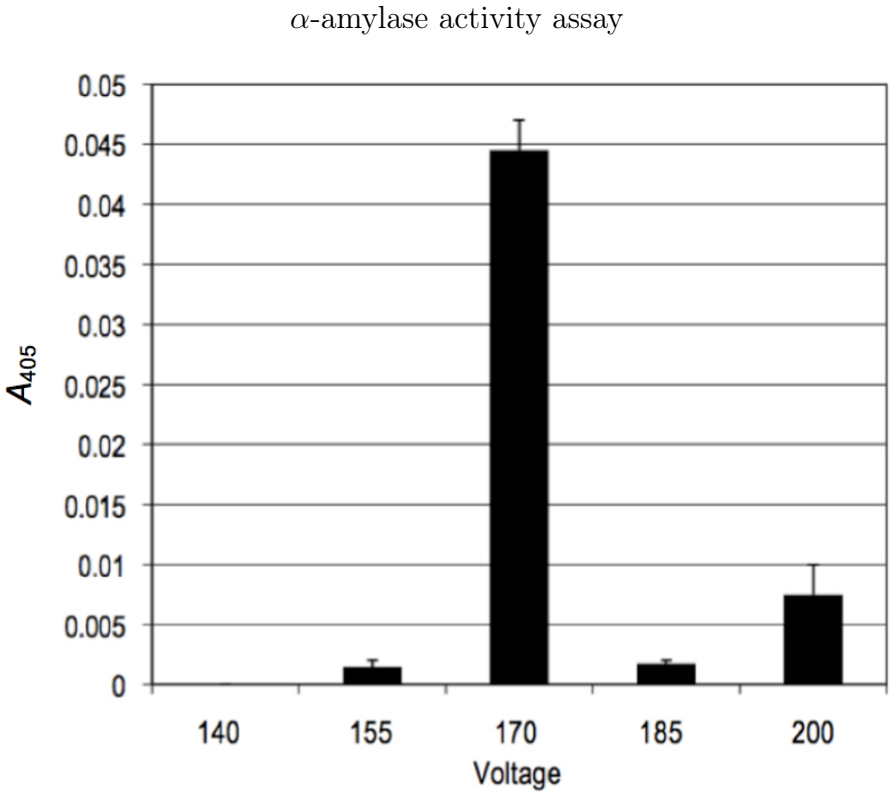

Figure 3.9: Activity of α-amylase was measured on *T. cuspidata* protoplasts using various electroporation conditions.

3.10 C] [Curtis and Grossniklaus, 2003]. The Gateway® cloning method [described in 2.4] utilises a site-specific bacteriophage lambda mediated recombination of DNA fragments [Earley et al., 2006]. An initial amplification of the DNA fragment of interest uses specific primer pairs carrying AttB overhangs necessary for recombination into the generic pENTRY vector during the BP reaction. The latter vector, harbouring a bacterial origin of replication is designed for recombination and propagation within the *E. coli* host. The subsequent LR reaction transfers the cloned fragment into the specific destination vector pMDC162 harbouring the Gateway® cassette, which drives the *GUS* reporter gene, and left and right border for *Agrobacterium* mediated expression in *N. benthamiana*.

The 7 week old *N. benthamiana* plants, past their 9 leaf developmental stage were used for infiltration. The *Agrobacterium* suspension carrying the chimeric *T. cuspidata TASY* and *DBTNBT* promoter *GUS* constructs were infiltrated in concert with the silencing suppressor construct p19 [Voinnet et al., 2003]. The promoter constructs and silencing suppressor were infiltrated at an optical density at a wavelength of 600 nm of OD 0.5 and OD 0.2 respectively. A volume of 100 μM of the *Agrobacterium* suspension resulted in an infiltrated leaf area of 3-4 cm^2 which was marked for ease of subsequent treatment and leaf disc incision. Treatment with MeJA was performed by spraying the plants with 200 μM MeJA in 0.2 % ethanol. The infiltrated and treated plants were incubated in a growth chamber at 22°C under 16h light. Four leaf discs with a radius of 8.5 mm were excised 30 h post infiltration and fluorometric *GUS* assays were performed to assess the MeJA-responsiveness

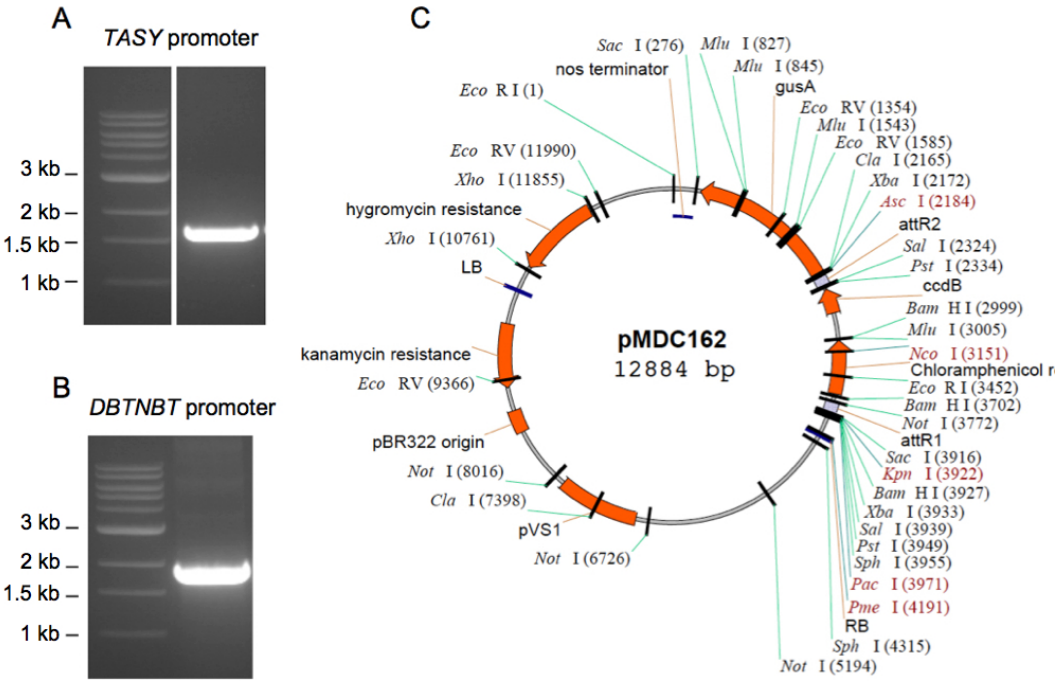

Figure 3.10: [A]: The amplified *TASY* and [B]: *DBTNBT* promoter bearing Gateway® the overhangs; [C]: The Gateway® vector pMDC162.

of the *TASY* and *DBTNBT* gene promoters. GUS activity was detected in infiltrated leaves with the *TASY* promoter driving the *GUS* reporter gene. However upon jasmonate treatment no GUS activity was measured using the *DBTNBT* promoter construct in *N. benthamiana* [Fig.: 3.11].

3.6 Discussion

The isolation and characterisation of promoter sequences is an important step en route to uncovering transcriptional regulation leading to gene expression [Terauchi and Kahl, 2000]. Attempts have been made to construct and screen

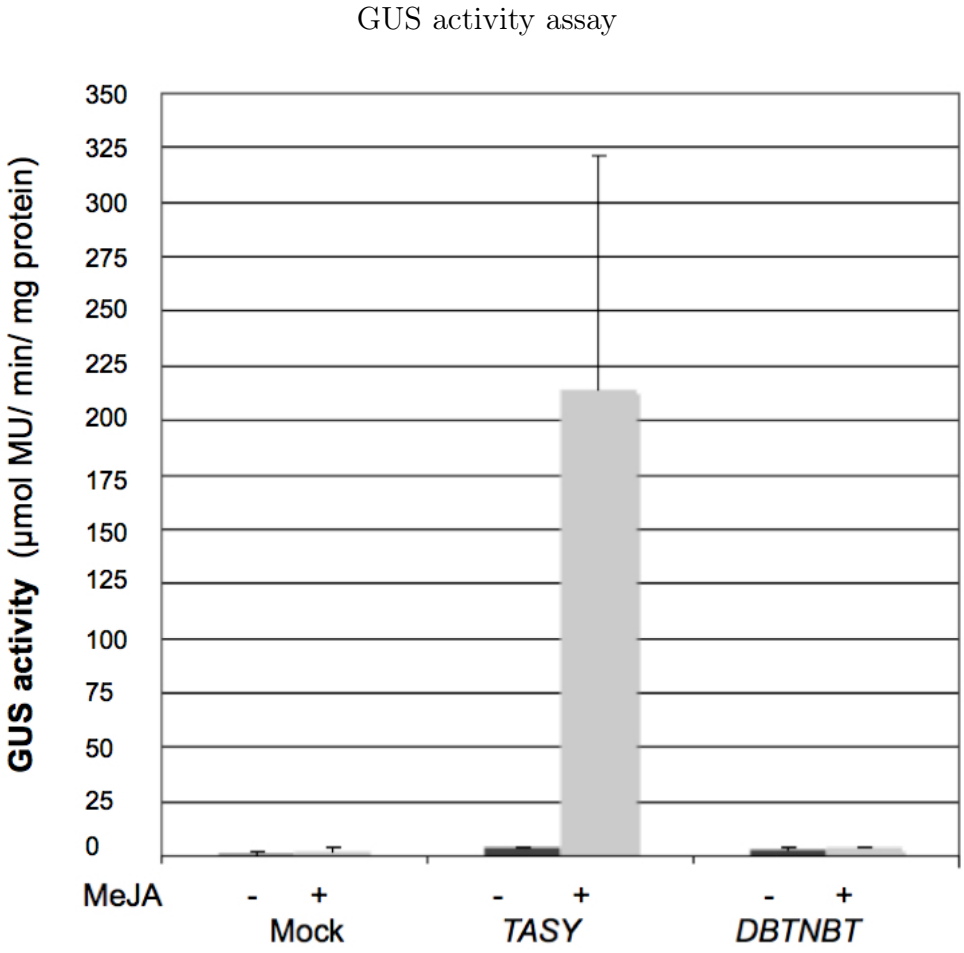

Figure 3.11: GUS activity assay using the *TASY* and *DBTNBT* promoters in the presence and absence of methyljasmonate. Mock: infiltration buffer only.

a genomic BAC library in order to isolate the promoter regions of the Taxol™ biosynthetic genes *TASY*, *BAPT* and *DBTNBT*. *T. cuspidata* is a diploid plant species with a haploid genome size of 10,850 Mb [Murray, 1998]. A discriminatory factor when constructing a BAC library is the probability of detecting the gene of interest. This probability can be determined by $P = 1\text{-}e^{N[ln(1-I/GS)]}$, [Farrar and Donnison, 2007]. Applying this calculation to the *T. cuspidata* genomic BAC library would require, for a probability of 98% to

detect a given gene, a number of 581,442 BAC clones [Appendix II]. However, the calculated probability assumes that the gene of interest is not toxic in the bacterial host and as such doesn't cause any lethal effect or rearrangement [Farrar and Donnison, 2007]. Furthermore empty BAC vectors or cfus without a BAC construct, as observed among the *T. cuspidata* genomic library, are not integrated in this equation, rendering the calculated probability a theoretical value.

Although two size selections were performed with a target size of 100-300 kb the average insert size of 73 kb was lower than expected. The transformation efficiency of the employed electrocompetent cells also showed a lower cfu frequency. A higher vector to insert ratio and high strength T4 DNA ligase did not result in increased transformation efficiency. The observed transformation efficiency of the utilised competent cells was $1x10^5$ cfu/μg 73 kb BAC insert which corresponds to about 1.2% of the expected outcome. These factors rendered the construction of a genomic BAC library from *T. cuspidata* impractical. However there are examples of BAC libraries of plant species with large genomes such as *Hordeum vulgare* (Barley) which has a haploid genome size of 5,000 Mb. The associated BAC library contains 313,344 BAC clones with an average insert size of 110 kb, which equates to a genome coverage of 6.3 [Yu et al., 2000]. Furthermore a method describing the storage and PCR based target gene isolation of a single BAC clone has been described using *Allium cepa* (onion) which also harbours a large haploid genome of 15,000 Mb [Suzuki et al., 2002]. The described storage and screening methodology could be applied to a *T. cuspidata* BAC library.

Successful isolation of a 513 bp fragment corresponding to the *TASY* promoter was performed using genome walking. Similar adapter ligation PCRs have been used frequently to isolate un-cloned genomic DNA fragments [Siebert et al., 1995]. Five different restriction enzymes were used in order to increase the likelihood of generating a genome-adapter fragment from partially digested genomic DNA whose length would conform to the applied PCR conditions. Although this approach was applied to amplify promoter fragments from *TASY*, *BAPT* and *DBTNBT*, sequencing of the amplified fragments was not able to confirm the authenticity of the *BAPT* and *DBTNBT* fragments. It is possible that identical sequences to the primer annealing sites of *BAPT* and *DBTNBT* within the *T. cuspidata* genome exist, which result in the amplification of specific yet not authentic PCR products.

The cloned 513 bp promoter fragment was further confirmed as genuine by comparison to a 1436 bp long sequence from the NCBI public domain corresponding to the *TASY* promoter.

In-silico analysis of the *TASY* and *DBTNBT* promoter revealed several jasmonate and pathogen-responsive *cis* elements. At position -142 within the proximal *TASY* promoter a GCC-box was located. A central GCC-box is part of the JERE element within the *STR* promoter in *C. roseus* [Menke et al., 1999]. JERE interacting TFs are AP2 TF family members and confer jasmonate regulated gene expression [van der Fits and Memelink, 2000].

The proximal *TASY* promoter harbours further jasmonate-responsive elements namely a G-box at position -158 and a related T/G-box at -255. Two G-boxes are also present in the *DBTNBT* promoter: one in the proximal

region at position -92 and a further in the upstream promoter region. G-box elements have been found to be over represented in promoters of jasmonate-responsive genes [Mahalingam et al., 2003]. The bHLH TF *Cr*MYC2 was found to interact with the G-box and T/G-box. Originally isolated using the G-box sequence (CACGTG) as bait in a yeast one-hybrid screen, *Cr*MYC2 showed interaction with the T/G-box element within the ORCA3 promoter in a jasmonate-responsive fashion [Pre et al., 2000].

A CGTCA motif is also located at position -189 with the proximal *TASY* promoter. In *Hordeum vulgare* the CGTCA motif exists as inverted repeats where it interacts with a TF of the bZIP family to confer jasmonate-responsiveness [Rouster et al., 1997].

Both analysed promoters were found to harbour several W-boxes of which two are located within the proximal regions at positions -136 and -92 in the *TASY* and *DBTNBT* promoter respectively. Pathogen-responsive gene expression has been shown to be regulated by WRKY TFs which bind to W-box elements [Rushton and Somssich, 1998]. It has also been shown that fungal elicitors contribute to Taxol^{TM} accumulation [Xu et al., 2004].

The MYB consensus motif CNGTTR is present twice in the *TASY* and three times in the *DBTNBT* promoter. This motif conforms to the MYB binding site I [Romero et al., 1998]. In *Petunia hybrida* the TF MYB.Ph3 which is involved in regulation of flavonoid biosynthesis recognises a GTTA motif, which forms part of the MYB I motif [Solano et al., 1995]. A further MYB II binding consensus motif was also located within the proximal promoter *DBTNBT* at position -261.

In order to functionally characterise the *TASY* and *DBTNBT* promoters, attempts have been made to develop a transient transformation assay utilising the *T. cuspidata* procambium csc. The advantage of effective transient assays is the relatively quick analysis of chimeric gene constructs rather than time consuming and laborious stable transformation [Hadlington and Denecke, 2001].

Particle bombardment was performed with a rupture disc pressure of 1350 psi in combination with 1.6 μm microcarriers with three replicate shots at a target distance of 3 cm which was found to achieve the highest transformation efficiency. However analysis of a *TASY* promoter *GUS* fusion construct revealed no detectable GUS activity. All transformations were performed using 0.5 μg of plasmid DNA and 500 μg of microacarriers. A similar study using particle bombardment utilised four different somatic csc from *Taxus* species, established varying transformation susceptibilities among the bombarded employed cell lines [Vongpaseuth et al., 2007].

Protoplast isolation and electroporation was performed using *T. cuspidata* procambium csc. As plant cell walls restrict the movement of macromolecules, protoplasts provide an attractive alternative for transformation [Bates, 1999]. A previous report demonstrated the isolation of protoplasts from *Taxus* for the purpose of protein content analysis by flow cytrometry [Naill and Roberts, 2005b]. However transformation of *Taxus* protoplast has not been shown.

To relate the employed enzymes used for protoplast isolation to the given cell wall composition, the effects of three different commercial cellu-

lase preparations were tested. cellulase hydrolyses 1,4-β-D-glucosidic linkages in cellulose, lichenin and cereal β-D-glucans [Buchholz et al., 1983]. Employed were cellulase-*Trichoderma*, cellulase Onozuka RS and cellulase Onozuka R10. cellulase-*Trichoderma* 1% [w/v] in combination with 0.1% [w/v] pectolyase, from *Basidiomycetes* species which exhibits endo-polygalacturonase and endo-pectin lyase activity, and 0.2% [w/v] Driselase^{TM}, an enzyme preparation from *Basidiomycetes* species which includes laminarinase, xylanase and cellulase activity was found to generate the highest observed yield of viable protoplasts.

Successful electroporation of plant protoplasts depends on various parameters including protoplast diameter and the utilised medium composition [Saunders et al., 1995]. To identify conditions suitable for the transformation of *T. cuspidata* protoplasts by electroporation, increased voltages were used starting from 140 V to 200 V in 15 V increments. To assess the electroporation efficiency the standardised α-amylase reporter assay was performed which revealed the highest activity at 170 Volts. The utilised TEX buffer is also used for the electroporation of *N. benthamiana* protoplasts with optimum electroporation conditions of 160 Volts [Hadlington and Denecke, 2001]. Although α-AMYLASE activity was recorded, the achieved activity relating to the transformation efficiency was 2.5 fold below the expected outcome rendering this approach insufficient as readings below 0.1 are unreliable [Prof. J. Denecke, personal communication].

To overcome the transformation limitations experienced with the model plant *T. cuspidata* an established approach of *Agrobacterium*-mediated transient expression in leaves of *N. benthamiana* was adopted [Yang et al., 2000].

This method targets more leaf cells thus increasing the transformation efficiency compared with particle bombardment [Levy et al., 2005]. The chimeric constructs harbouring the *GUS* reporter gene driven by the *TASY* and *DBTNBT* gene promoters were introduced into *N. benthamiana* leaves and treated with methyljasmonate. *GUS* expression was observed by the *TASY* construct, however, no GUS activity was seen in the *DBTNBT* construct. The samples were analysed 24 h after treatment at which point expression of both genes in the *T. cuspidata* csc was found to be high by RT-PCR [chapter 5].

Located within the proximal *TASY* promoter are several putative jasmonate-responsive elements including a GCC-box, a CGTCA motif and a G- and T/G-box. These conserved *cis* elements provide binding sites for jasmonate-responsive TFs. Although a G-box is also present within the *DBTNBT* promoter, no jasmonate-responsiveness was observed after treatment. This may indicate that *DBTNBT* expression in *T. cuspidata* is controlled by a more complex transcriptional regulation not conserved in *N. benthamiana*.

Chapter 4

Transcriptome sequencing of *Taxus cuspidata*

The employed csc was initiated from isolated procambium cells of *T. cuspidata* [Lee et al., 2010]. The procambium, which gives rise to the vascular tissue, is formed by the apical meristem and can be regarded as vascular stem cells [Fukuda, 2004]. Among the unique characteristics of the procambium derived csc, in this thesis referred to as cambial meristematic cells [CMCs], is an increased production of Taxol™ of 443 % with 102 mg/kg FCW. In contrast a csc initiated from somatic dedifferentiated needle cells [DDCs] in a 125 ml culture volume accumulated 23 mg of Taxol™ /kg FCW [Lee et al., 2010]. To facilitate the identification of jasmonate-responsive TFs as potential regulators, which govern transcriptional control of the Taxol™ biosynthetic pathway and may be responsible for the observed increase in Taxol™ production, the sequencing of the *T. cuspidata* transcriptome was performed. A

further rationale was the identification of genes differentially expressed in the procambium csc when compared to the somatic csc, which may be associated with its stem cell identity.

4.1 Sequencing Methods

As *T. cuspidata* is not a widely used model organism, only limited sequence data resources were available at the onset of this project. However, non-model organisms especially possess valuable traits, which emphasises the need for sequence data resources necessary for functional molecular studies [Coemans et al., 2005]. These traits can be important pharmaceutically [Taxol^{TM}: *T. cuspidata*], economically [Banana crop: *Musa acuminata*], ecologically [Pine foundation species: *Pinus contorta*] or from a population biology perspective [Glanville fritillary butterfly: *Melitaea cinxia*] [Coemans et al., 2005, Parchman et al., 2010, Vera et al., 2008].

Historically EST data has provided an attractive method for obtaining sequence data from non-model organisms where whole genome data is unavailable. This approach is especially appealing with organisms harbouring a large genome, such as Pine, as ESTs are representative of functional open reading frames, lacking introns and intragenic regions [Parchman et al., 2010]. Recent advances in next-generation sequencing technologies such as Roche/ 454 and Illumina/ Solexa provide cost-effective, fast and highly accurate approaches for the generation of sequencing data [Mardis, 2008]. With the focus on the transcribed portion of the genome, the generated cDNA used in transcriptome

sequencing minimises the target sequencing size and thus reduces input time and labour. A further advantage of the combination of Roche/ 454 and Illumina/ Solexa sequencing technologies is the versatility of the approach. Whilst Roche/ 454 allows for the de-novo sequencing of the transcribed genes in *T. cuspidata*, Illumina/ Solexa *Nla*III tag sequencing will monitor the expression profile thus providing qualitative and quantitative data respectively [Mardis, 2008], 2008; www.illumina.com).

4.1.1 The Roche/ 454 technology

The Roche/ 454 Life Science sequencing technology [454, Bradford, USA; since 2007 Roche, Basel, Switzerland ; http://www.454.com] was introduced to the scientific community in 2004 [Rothberg and Leamon, 2008]. The introductory paper from 2005 is cited to date over 600 times which demonstrates the immense impact of this technology [Margulies et al., 2005]. Ambitious projects employing 454 sequencing include the de-novo sequencing of bacterial genomes [Margulies et al., 2005], the genome sequencing of an individual human [Wheeler et al., 2008] and the partial sequencing of a Neanderthal genome [Briggs et al., 2007, Green et al., 2006, Noonan et al., 2006].

The strength of Roche/ 454 is the generation of long sequencing reads with the current technology achieving up to 400 bp and a total of 1 X 10^6 sequence reads [Voelkerding et al., 2009]. The Roche/ 454 sequencing technology employs chemiluminesence coupled with pyrosequencing [Fig.: 4.1]. In the initial emulsion PCR, designed to amplify the target DNA, sheared fragments of 400-600 bp in length which are attached to streptavidin beads through

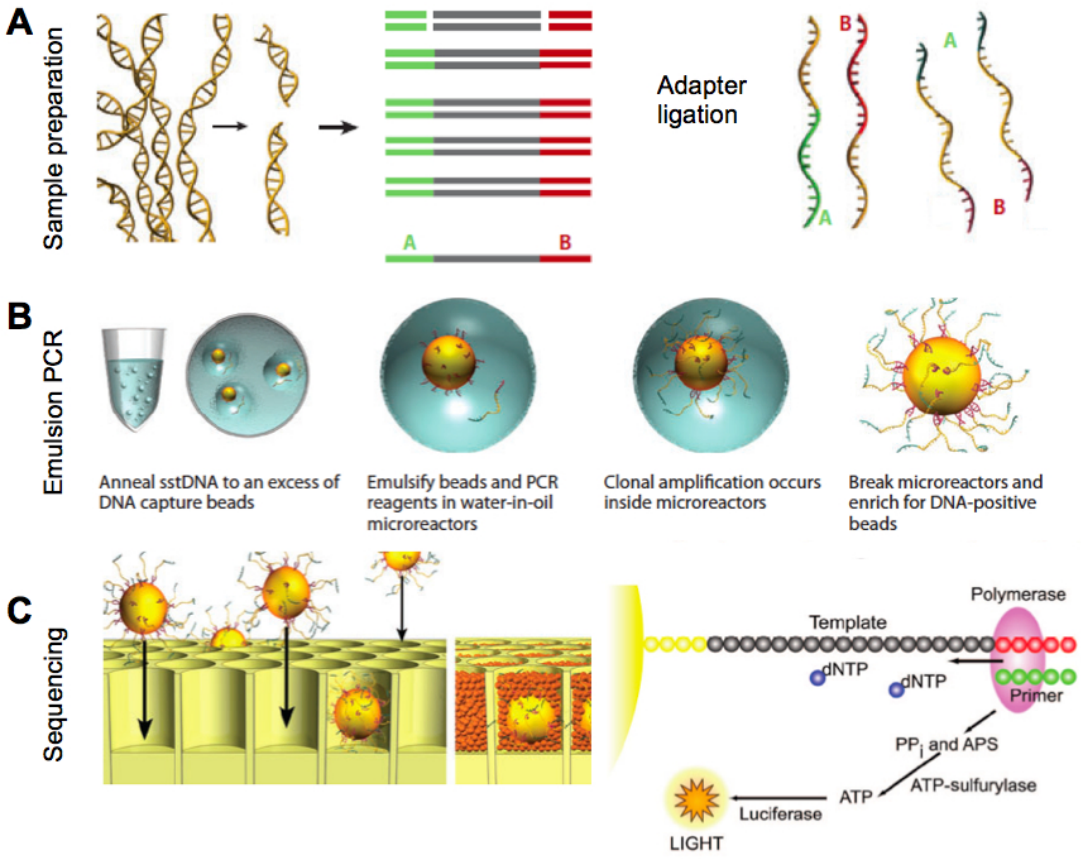

Figure 4.1: Diagram showing the Roche/ 454 methodology. [A]: Sample preparation: the sheared single stranded (sst) DNA fragments are ligated to specific adapters; [B]: Emulsion PCR: the sst DNA fragments are annealed to capture beads within emulsion droplet microreactors for clonal amplification; [C]: Sequencing: each individual bead is placed into a single well of the picotiter plate. Enzymes and reagents necessary for pyrosequencing are added to the beads and the primer is hybridized to the template. Upon incorporation of one of the sequentially added dNTPs, ATP sulfurylase catalyses the conversion of the released pyrophosphate (PP_i) and adenosine 5'-phosphosulfate (APS) into ATP and sulphate. The ATP is utilised in a subsequent reaction with luciferin in which luciferase catalyses their conversion into oxyluciferin and visible light. Source: [Mardis, 2008, Voelkerding et al., 2009]

specific adapters are used. Each individual fragment is captured in an emulsion droplet which provides the amplification environment resulting in 1 X 10^7 clonal copies of the single-stranded target DNA [Margulies et al., 2005]. A subsequent pyrosequencing reaction using chemiluminesence is carried out where each target bead is transferred into a well of a picotiter plate thus allowing parallel reactions. In this sequencing-by-synthesis approach, nucleotides are added into the wells in a sequential fashion, causing the release of PP_i through incorporation of complementary nucleotides to the target DNA fragment. The light emitted through released PP_i is monitored and nucleotides of target DNA determined [Morozova and Marra, 2008, Mardis, 2008].

4.1.2 The Illumina/ Solexa technology

The Illumina/ Solexa [San Diego, USA.; http://www.illumina.com] sequencing approach concentrated on the generation of expression profiles for the *Taxus cuspidata* transcriptome generated by the Roche/ 454 technology. Although generated sequence reads are only 25-35 bp in length, the employed *Nla*III tag sequencing methodology creates a digital gene expression profile by mirroring the frequency of a sequenced tag through its occurrence in the cDNA population [Mardis, 2008] [Illumina work flow sheet].

During the initial stages of sample preparation, a signature tag corresponding to the 3' end of each transcript is generated using the *Nla*III restriction endonuclease [Illumina protocol]. Following a bridge amplification step, which produces 1 X 10^7 clonal copies of the target fragment, the sequence is generated in a sequencing-by-synthesis approach. DNA polymerase and all

four nucleotides are provided simultaneously. The latter each carry a specific fluorescent label, which is monitored during incorporation into the newly synthesised signature tag [Mardis, 2008, Voelkerding et al., 2009].

4.2 Experimental design

The csc of DDCs and CMCs from *T. cuspidata* were employed for the sequencing of the transcriptome. We reasoned that a comparison of gene expression profiles present in the CMCs to DDCs would reveal genes preferentially expressed in the procambium thus contributing to its given meristematic character. To accommodate the identification of TFs a set of jasmonate treated samples were employed. Both cell type cultures of the *T. cuspidata* [5 days after subculturing] were adjusted with fresh medium to 100 ml and grown for 2 h in the dark to acclimatise. The cultures were elicited with 100 μM MeJA and samples were taken at 0.5 h, 2 h and 12 h post elicitation [Fig.: 4.2]. The total RNA was extracted using a RNesay plant RNA kit [Qiagen, Venlo, Netherlands] following the manufacturer's instructions from all samples and the integrity of the RNA was subsequently assessed by agarose gel electroporation. To verify the MeJA treatment of the samples, RT-PCR was performed showing the up-regulation of two Taxol^{TM} biosynthetic genes, *TASY* and *DBBT* in response to jasmonate treatment [Fig.: 4.3]. To enable subsequent statistical analysis three biological replicates were generated. The sample used for Roche/ 454 sequencing was prepared using equal amounts of the total RNA designated for Illumina/ Solexa sequencing.

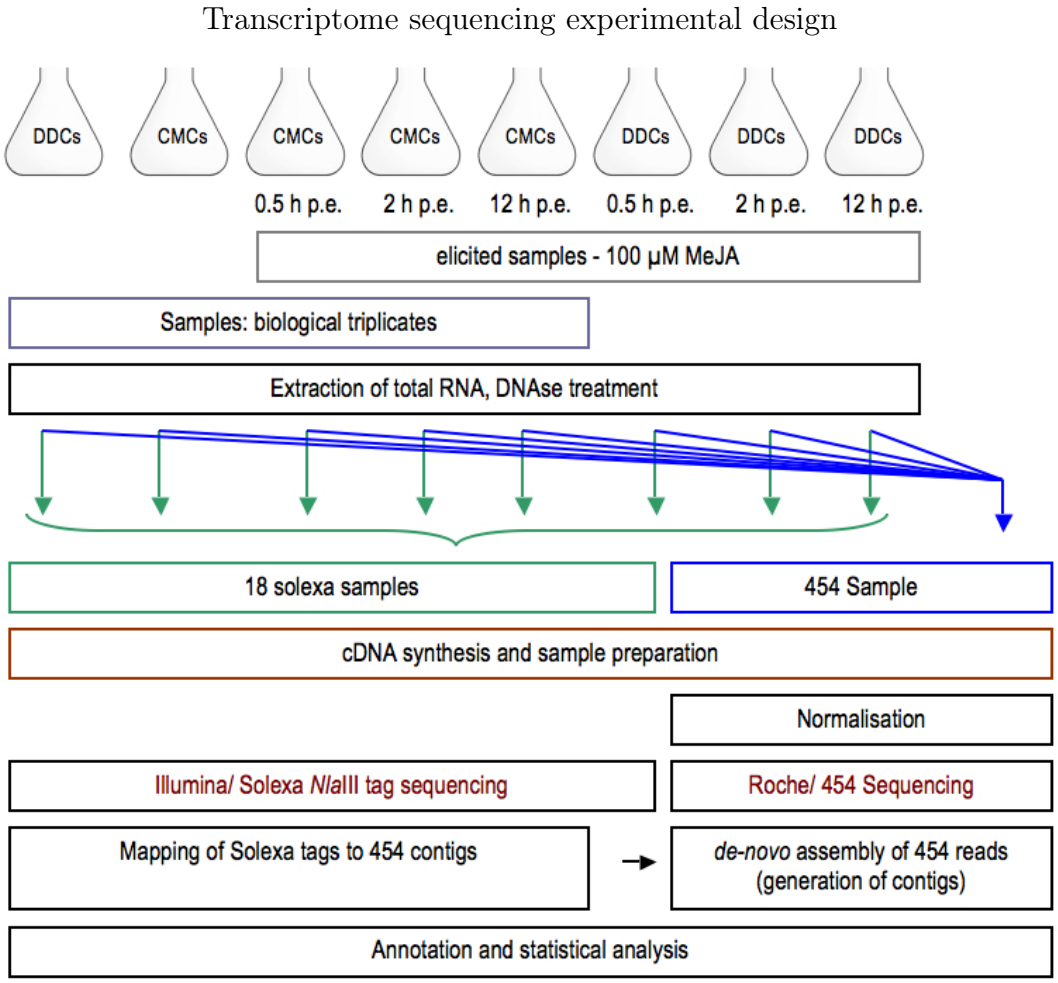

Figure 4.2: Schematic showing the experimental design including various samples, RNA extraction, sample preparation, sequencing, assembly and annotation.

Note that the subsequent steps including DNAse treatment, cDNA synthesis, normalisation of the 454 sample using Kamchatka crab duplex-specific endonuclease [Zhulidov et al., 2004], Roche/ 454 and Illumina/ Solexa sequencing, mapping, annotation of the generated sequence contigs were performed by the Gene-pool genomics facility at the University of Edinburgh and the

statistical analysis using EdgeR [Robinson et al., 2010] was done by Florian Halbritter.

MeJA elicitation of transcriptome samples

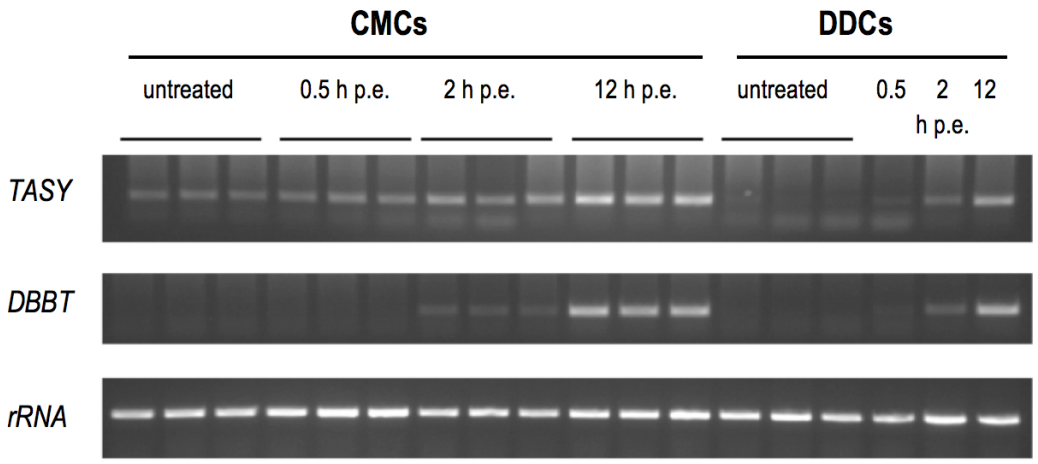

Figure 4.3: RT-PCR showing the induction of Taxol^{TM} biosynthetic genes *TASY* and *DBBT* after elicitation with 100μM MeJA in transcriptome sequencing samples prepared from the CMCs and DDCs. rRNA: ribosomal RNA.

4.3 Transcriptome characterisation

Roche/454 sequencing was performed on a Roche GS FLX genome sequencer. The sample prepared from *T. cuspidata* cDNA was quantified and 2,000,000 beads were loaded onto a GS FLX pico titre plate. Roche/ 454 pyrosequencing achieved a total of 860,800 reads with an average length of 351 bp, thus a total of 302 Mbp was generated [Table: 4.1].

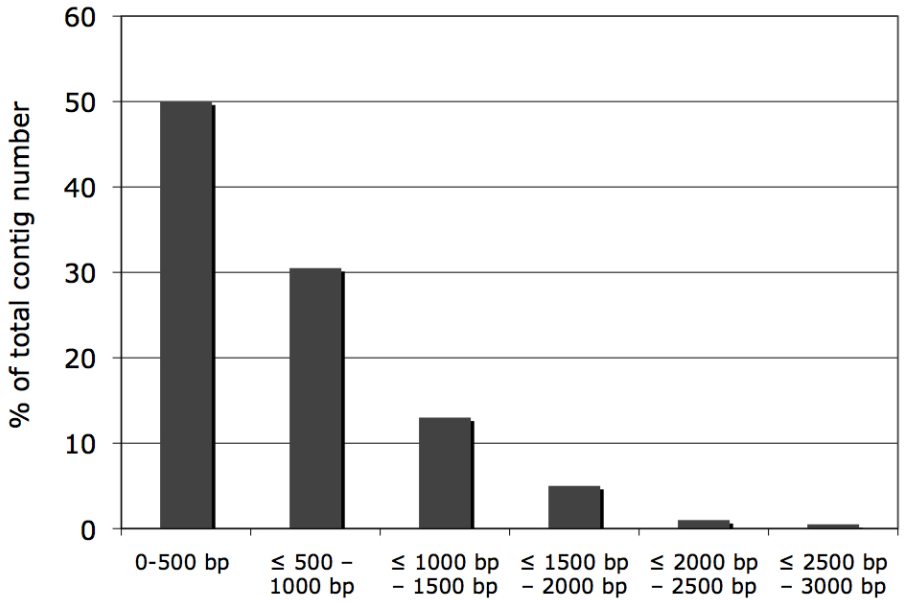

Figure 4.4: Contig length generated by Roche/ 454 sequencing and subsequent assembly using Newbler Assembler 2.0.1 as percentage of the total transcript count of 36,823.

The 454 sequence reads were assembled into contigs using the Newbler Assembler 2.0.1 [454 Life Science] software, which generated a total of 36,906 contigs with an average length of 699 bp. Comparison of all Kamchatka normalised contigs revealed that 30,823 [83.5% of the total contig number] were unique within the *T. cuspidata* transcriptome [Table: 4.1]. The largest group of contigs [50%] are up to 500 bp in length. However, a substantial number [31%] are in the size group of 500 bp - 1500 bp and 6.5% of the contigs are 1500 bp - 3000 bp in length. The maximum contig length within the transcriptome is 10355 bp [Fig.: 4.4].

A few months later a second assembly was performed using the new Newbler Assembler software version 2.3 [454 Life Science]. Although both programs are specifically designed for the assembly of Roche/ 454 pyrosequencing reads,

Assembly and annotation statistics

Roche/ 454	
total number of reads	860,8001
average length	351 bp
Newbler Assembler 2.0.1	
total number of contigs	36,905
unique contigs	30,823
average contig length	699 bp
Newbler Assembler 2.3	
total number of isotigs	19,614
Annotation	
total number of annotated contigs	18,173
contigs with uncharacterised hit	12,902
contigs with characterised hit	5,271
contigs with no annotation	12,650

Table 4.1: Transcriptome statistics of Roche/ 454; number of reads, assembly and annotation using two different versions of the Newbler Assembler and subsequent annotation.

the Newbler Assembler 2.0.1 was originally designed for linear assembly of genomic data. In contrast the Newbler Assembler 2.3 supports incremental assembly thus taking alternative splicing events into account. This assembly generated 19,614 isotigs. In summary, the assemblies performed using the Newbler Assembler 2.0.1 and 2.3 mirror the transcript and expressed gene content in *T. cuspidata*. Note that all subsequent analysis of the transcriptome data was performed using the contig assembly of the Newbler 2.0.1 version.

The contigs were subject to BLAST searches [http://blast.ncbi.nlm.nih.gov/Blast.cgi] using the plant protein entries using the Uniref100 [UniProt Reference Clusters] database using an E-value cut-off of e 1 x 10^{-3}. This generated 18,173 annotated contigs [59% of the total contig number] indicating

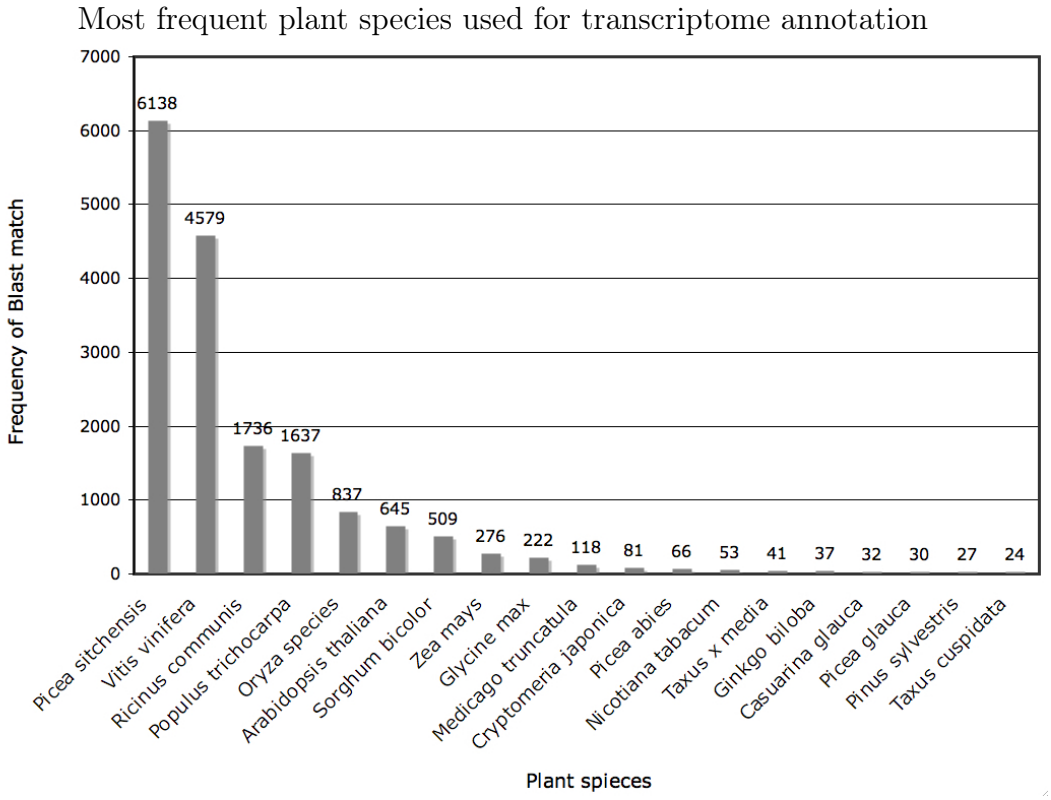

Figure 4.5: The 19 top plant species used for the annotation of the transcriptome. This includes the coniferus tree *Picea sitchensis*, *Taxus media* and *Taxus cuspidata*.

that *T. cuspidata* has a vast gene content of 12,650 transcripts which are unique to the genus *Taxus*. However, of the annotated contigs 12,902 or 71% show the highest similarity to a gene with an unknown function. The remaining 29% corresponding to 5271 contigs have a protein function assigned [Table: 4.1]. The contigs from *T. cuspidata* showed significant similarities to entries within 60 different plant species used for annotation. It is noteworthy that within the annotated fraction of 18,173 contigs 34% showed the highest similarity to *Picea sitchensis* [Fig.: 4.5]. Other plant species which show high BLAST matches

contig	*P. sitchensis*	*A. thaliana*	*O. sativa*	*Z. mays*	gene function
00317	76	70	64	42	abhydrolase 3
01434	69	49	64	53	DNA-binding protein
01458	67	50	52	49	alcohol dehydrogenase
00295	86	77	63	56	ccr4-associated factor
00541	76	68	73	71	peroxisomal biog. factor
00580	66	42	48	49	bZIP TF
00895	81	59	54	54	N-acetyltransferase
01276	74	47	67	n/a	glycosyltransferase
01730	66	50	51	46	zinc finger protein
03246	70	67	66	68	indole-3-glycerol phosphate synthase
	73%	**58%**	**60%**	**54%**	

Table 4.2: Sequence identities in % of *T. cuspidata* to other gymnosperm and angiosperm model organisms.

include *Vitis vinifera*, *Ricinus communis* and *Populus trichocarpa*.

To assess whether the given divergence time of *T. cuspidata* to other gymnosperm and angiosperm model organisms is reflected at the sequence level, identity comparison was performed. To this end the sequence of 10 randomly selected contigs from *T. cuspidata* which were annotated with homologues from *Picea sitchensis* were retrieved and the average amino acid sequence identity was scored. *P. sitchensis* showed a sequence identity of 73 % to the homologues *T. cuspidata* contig, *A. thaliana* 58 %, *O. sativa* 60 % and *Z. mays* 54 % [Table: 4.2]. The observed percentages underpin the expected sequence diversity which is reflected by the divergence time of *T. cuspidata* to other gymnosperm and angiosperm plant models.

A gene ontology classification was performed using Annot8r [Schmid and

Blaxter, 2008] which resulted in a significant portion being associated with a GO term. A total of 1139 GO terms which account for a wide functionality of genes were assigned to 7739 contigs. As the cDNA for Roche/ 454 sequencing was prepared using total RNA from both cell types and MeJA treated csc, the frequency of the GO terms are not representative of a single sample. In order to quantify the generated Roche/ 454 transcriptome data

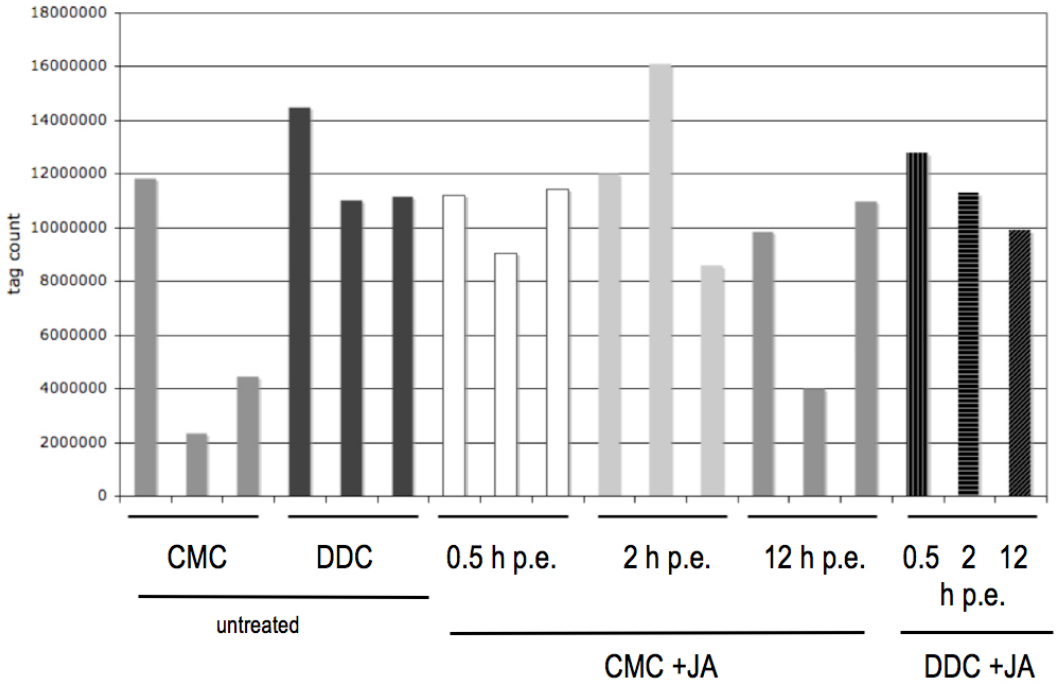

Figure 4.6: Achieved tag counts by Illumina/ Solexa sequencing across all samples from CMCs and DDCs.

Illumina/ Solexa *Nla*III DGE tag sequencing was carried out. The DGE approach combines the generation of a global gene expression profile with the detection of rare transcripts thus providing an ideal technology for non-model organisms ['t Hoen et al., 2008]. Sequencing was accomplished on a GAII

and GAII$_x$ Illumina sequencer. As two different Illumina Genome Analysers were employed the achieved number of tags clearly mirrors the performance of the next generation GAII$_x$ versus the GAII. Samples sequenced on the latter, namely CMC#2, CMC#3 and CMC#2-12 hours p.e. show a consistently lower tag count. DGE tag counts from samples sequenced on GAII$_x$ range from 8,584,203 to 16,102,621 [Fig.: 4.6]. The resulting expression values were normalised to transcript counts per million [TPM], thereby revealing expression levels ranging from 0 to 47914 DGE tag counts.

4.4 The procambium csc specific transcriptome

In order to characterise the expression profile present in the CMCs, the Illumina/ Solexa data for the untreated CMC and DDC samples was statistically analysed. The software used for the determination of differentially expressed contigs within the DGE data was the edge$R$ Bioconductor package [Robinson et al., 2010]. To facilitate a stringent analysis further statistical parameters were applied to filter out transcripts showing a less convincing differential expression profile. These included a false discovery rate [p-value] of 0.05 and a minimal difference in tag count of 10 across all samples [including replicates]. This approach identified a total of 563 differentially expressed contigs in CMCs compared to DDCs [Fig.: 4.7]. Of these 296 and 267 showed up- and down-regulated expression in CMCs respectively.

As previously described the annotation of the *T. cuspidata* transcriptome data was insufficient, thus careful annotation of the d.e. contigs was

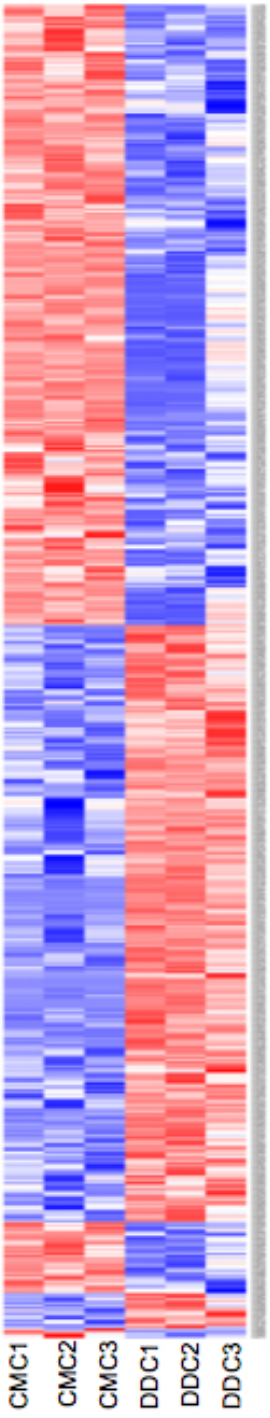

Figure 4.7: The heatmap shows the 563 differentially expressed contigs in the CMCs when compared to the DDCs. Contigs are clustered in rows with red and blue indicating up- and down-regulated expression values respectively. The sub-clusters at the bottom correspond to contigs which have an overall higher expression level.

performed. To this end a tBlastn analysis was done in which the translated nucleotide sequence is used to retrieve an homologue of the query sequence through which 375 contigs could be annotated. Subsequent determination of the corresponding GO term was performed computationally by the Gene-pool genomics facility at the University of Edinburgh. A total of 204 contigs could be associated with a term within 55 GO groups. The most frequent GO terms were grouped together and the presence of relative up- and down-regulated contigs was scored. This revealed that the procambium derived CMCs showed a noticable up-regulation of genes involved in "response to stress" and "defence response" [Fig.: 4.8]. Genes involved in lipid metabolism showed a significant

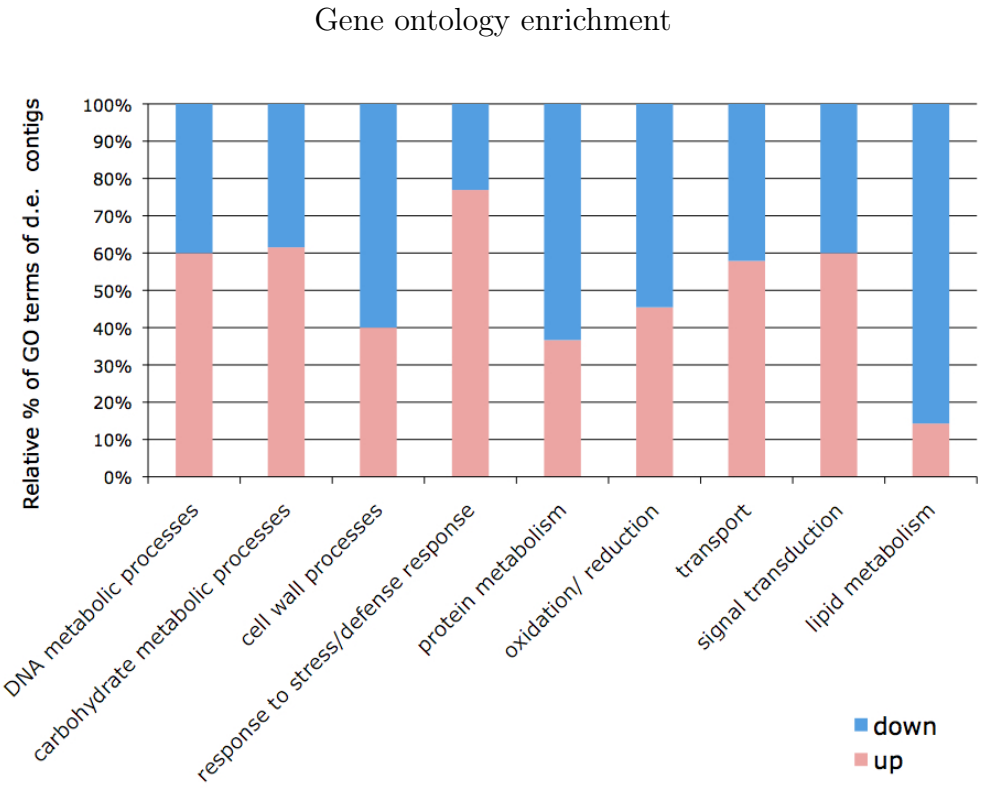

Figure 4.8: Most frequently occuring GO terms within the up- and down-regulated fraction among the 563 differentially expressed contigs in CMCs

down-regulation in CMCs. Furthermore, genes associated with DNA- and carbohydrate metabolic processes, transport and signal transduction showed a slight increase.

Sequence alignment of contig 01805 with *A. thaliana* and poplar homologue

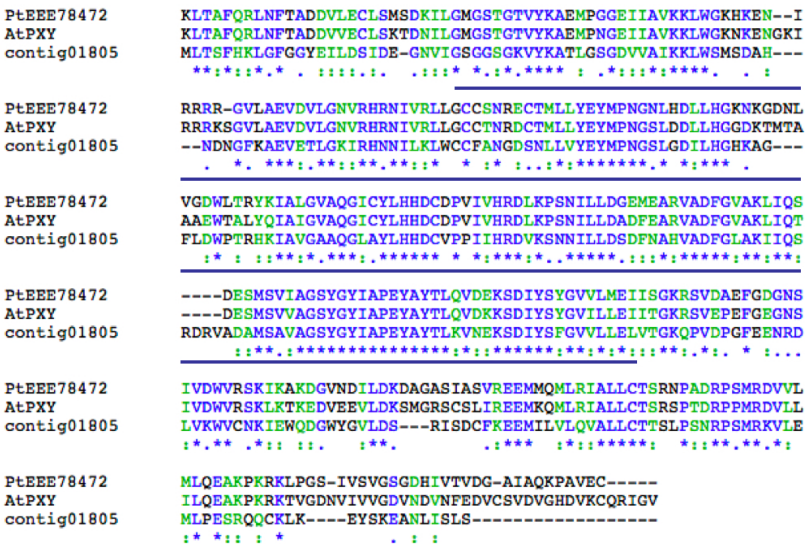

Figure 4.9: The multiple amino acid alignment shows the sequence similarity of the *T. cuspidata* clavata-like receptor with the authentic *PXY* (At5G61480) and the *Populus trichocarpa* homologue (EEE78472). The kinase domain is underlined in blue.

Within the d.e. genes in the CMCs is the clavata-like receptor contig 01805 which shows high amino acid similarity to the *A. thaliana PXY* gene and the *Populus trichocarpa* homologue [EEE78472] [Fig.: 4.9]. Members of this family exhibit a protein kinase domain. The *A. thaliana PXY* transcript is procambium localised and essential for ordered cell division and formation of vascular tissue [Etchells and Turner, 2010]. The expression profile of the *T. cuspidata* clavata-like receptor shows an increased (log2) fold-change of 2.1 in CMCs [Fig.: 4.10 A].

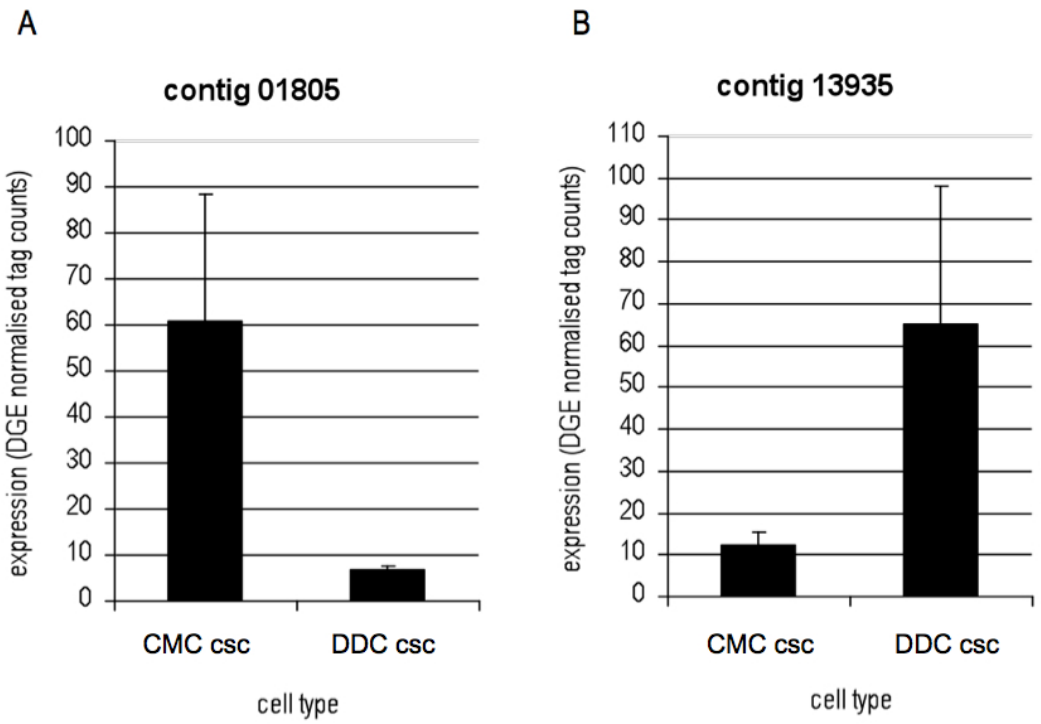

Figure 4.10: DGE expression data. [A]: The *T. cuspidata* homologue contig 01805 of the *A. thaliana PXY* gene; [B]: contig 13935 the AIF3 [AT3G17100] homologue.

Eight TFs were annotated among the 563 d.e. contigs in the CMCs of which through sequence analysis five *A. thaliana* homologues could be identified. Three TFs show up- and four down-regulation. Subsequent attempts to identify associated transcriptional networks using TAIR were futile. However, among the TFs is At317100, a bHLH TF [Appendix III] whose expression profile in *T. cuspidata* shows significant down-regulation in CMCs with a (log2) fold-change of -2.3 [Fig.: 4.10 B]. At317100 has recently been shown to interact with AtBS1 [Wang et al., 2009]. The latter is a component of brassinosteroid signalling which is involved in vascular bundle formation and early procambial

division [Ibanes et al., 2009].

Prior to the transcriptome sequencing, attempts had been made to confirm the procambium character of the employed csc by monitoring the expression of the *TED2* homologue in *T. cuspidata*. *TED2* is cambium specific and is involved in formation of tracheary elements, which are part of the plant conductive tissue [Fukuda, 1997].

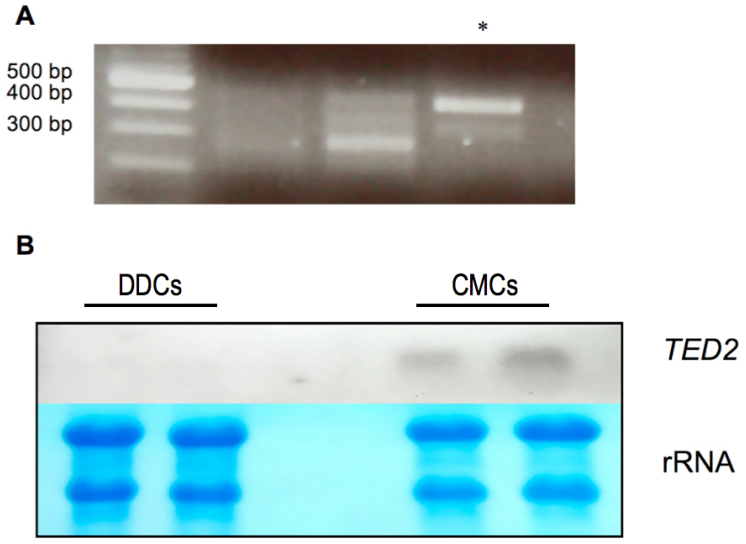

Figure 4.11: [A]: Degenerate PCR for the synthesis of the *TED2* northern probe, the asterisk indicates the 412 bp amplifed fragment; [B]: Northern analysis of the TED2 gene in the procambium [CMC] and somatic [DDC] derived csc. rRNA: ribosomal RNA.

The probe utilised for the *TED2* homologue in *T. cuspidata* was synthesised by degenerate PCR from CMC cDNA. This method allows for the amplification of DNA where no sequence information is available. Primer design was performed on the basis of multiple amino acid sequence alignments of the authentic *Zinnea elegans TED2* gene and two poplar homologues [Apendix I]. Regions which show a high degree of conservation were chosen for primer

design. The designed primers are the translated nucleotide representative of the amino acids and take sequence variations into account thus constituting a pool of very similar yet different specific primer combinations. Degenerate PCR generated a 412 bp long fragment [Fig.: 4.11 A] which showed a 69% sequence identity and 93% similarity to the authentic *TED2* fragment from *Zinnia elegans*. RNA from CMCs and DDCs was prepared and northern analysis was performed using the synthesised *TED2* fragment as a probe. A single band could be detected in the CMCs [Fig.: 4.11 B]. In contrast, no *TED2* transcript was detected in DDCs.

4.5 The JA-responsive transcriptome and *T. cuspidata* TFs

Transcription factors have a crucial function in regulating gene expression acting as mediators of signal transduction and transcription [Endt et al., 2002]. Jasmonates induce the biosynthesis of secondary metabolites in several plant species [Memelink et al., 2001, van der Fits and Memelink, 2000]. More importantly Taxol™ biosynthesis is jasmonate-responsive [Nims et al., 2006] thus identification of jasmonate inducible TFs provides a major step in elucidating the underlying transcriptional regulation in *T. cuspidata*.

Statistical analysis was carried out to compare the jasmonate-treated CMC samples over time with the assumption that contigs exhibit d.e. in at least one of the three employed time points, namely 0.5 h, 2 h and 12 h post elicitation and conform to a linear up- or down-regulation. This resulted in the

identification of 1646 d.e. contigs with a false discovery rate of 0.05 in CMCs.

In contrast analysis of the expression profiles in the jasmonate elicited DDC samples over time resulted in the identification of only 487 d.e. contigs. Surprisingly only 179 contigs showed d.e. in both cell types of which 89 contigs could be mapped to a GO term representing 40 different cellular processes. Comparing the GO term frequency and associated expression profiles within the two cell types after jasmonate treatment did not generate a coherent result. However, components of the jasmonate signalling pathway in *T. cuspidata* could be identified among the 179 shared d.e. contigs. GO term 0009867: "jasmonic acid mediated signalling pathway" was assigned to contigs 11289, 33381 and 22910 and GO term 0009753: "response to jasmonic acid stimulus" was assigned to contig 14280. Where as contig14280 is down-regulated in CMCs, contigs 11289, 33381 and 22910 are up-regulated in both cell types. Sequence analysis was able to verify these contigs as *T. cuspidata* homologues of the *A. thaliana* JAZ10, JAZ3 and JAZ2 and the *N. tabacum* JAZ3 repressors [Appendix IV]. The highest (log2) fold-change was observed in the 2 hour time point after jasmonate treatment which might coincide with active *JAZ* transcription mediated by MYC2 [Chini et al., 2007].

Hierarchical clustering analysis was performed using the data set of the 1646 d.e. contigs in CMCs visualised as a dendrogram and heatmap [Fig.: 4.12]. This exhibits the sample integrity of the biological triplicates which map closest to each other. However CMC#2-2 hours p.e. does not conform to the observed trend. The d.e. contigs show up- and down-regulation in response to jasmonate treatment. The transcriptional changes mediated by jasmonate

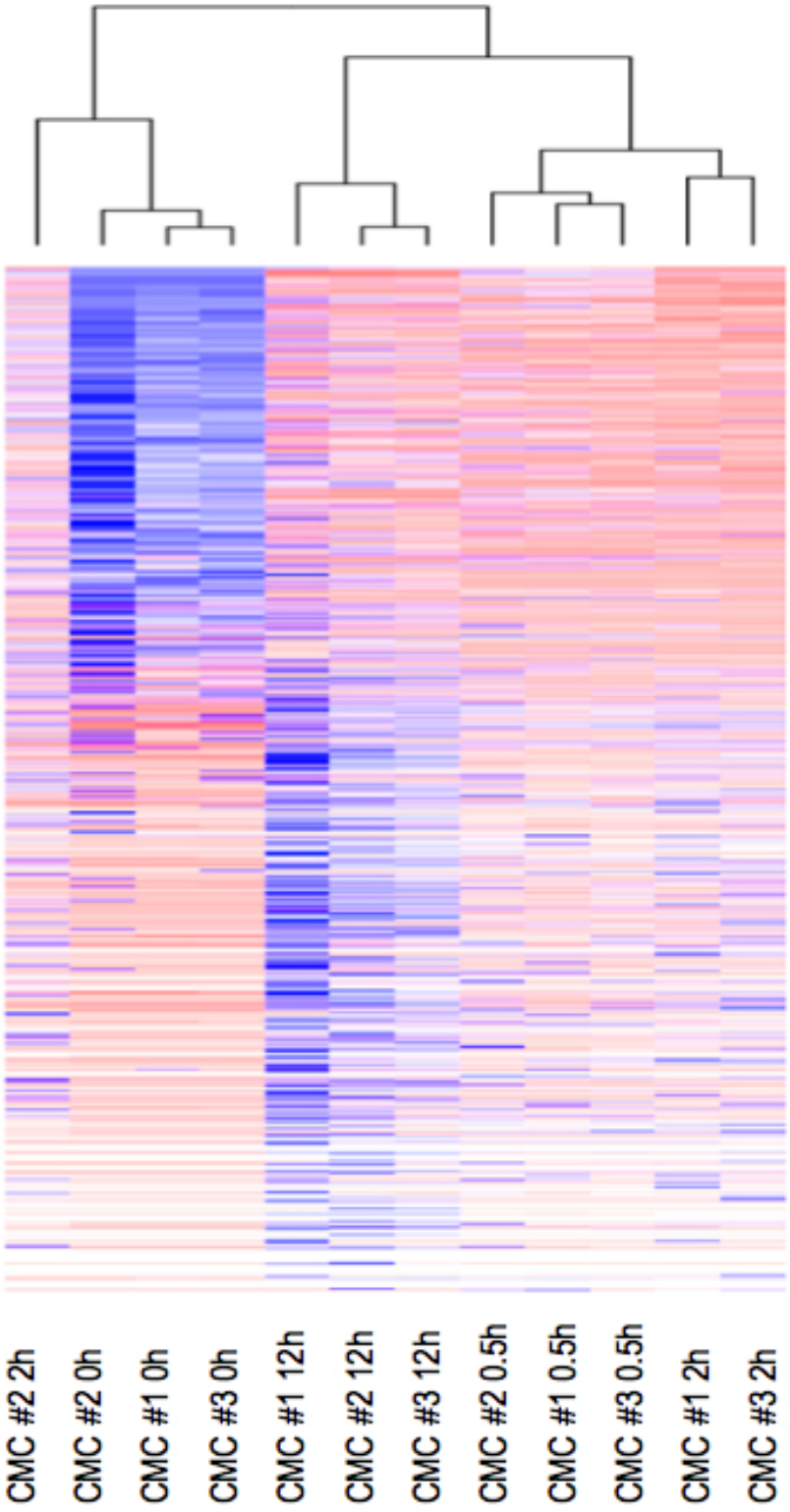

Figure 4.12: Heatmap and dendrogram of the 1646 d.e. contigs in at least one time point in response to jasmonate treatment.

is demonstrated by the dendrogram position of the 12 hour samples which maps closest to the untreated population followed by the 0.5 hour samples.

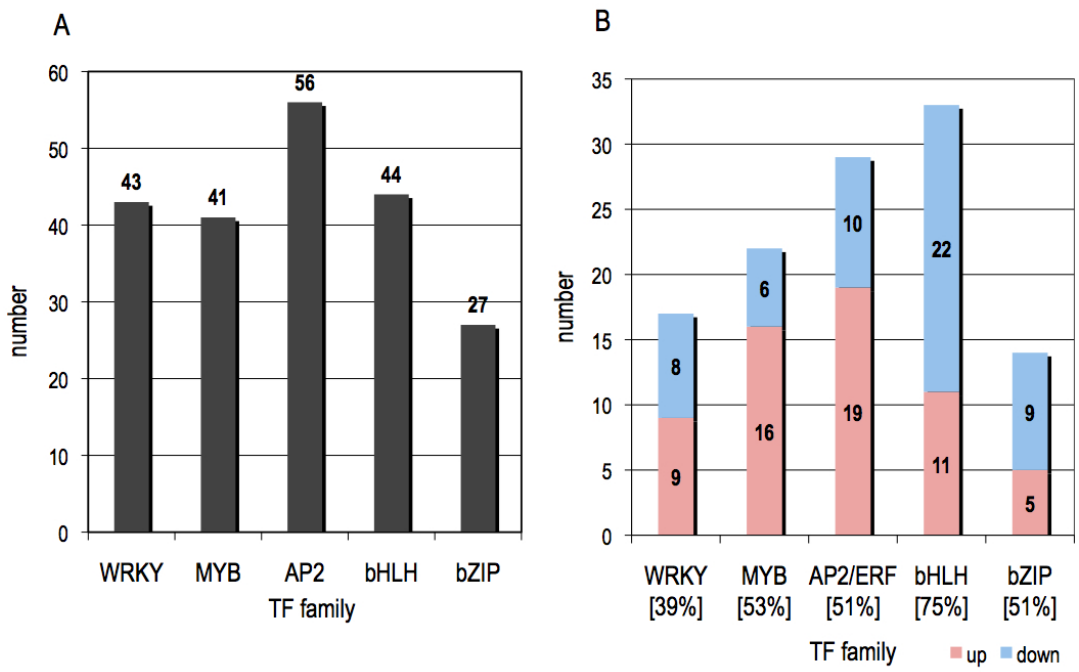

Figure 4.13: Identified TFs in *T. cuspidata*; [A]: total numbers of identified TFs in *T. cuspidata* from five different families; [B]: number of up- and down-regulated TFs in response to jasmonate treatment in at least one time point CMCs. [%] represent the number of jasmonate-responsive TFs relative to all family members.

As previously described, annotation resulted in only 17 % of the total contig number being assigned to a protein with known function. Thus Blastx [NCBI] searches were performed using amino acid queries from different poplar and *A. thaliana* TFs. The latter were retrieved from the Plant Transcription Factor database. The identification concentrated on TF families which have been shown to be JA-responsive and known to function as regu-

lators of secondary metabolism namely AP2/ER, MYB, bHLH, WRKY and bZIP TFs [Pauwels et al., 2008]. The putative *T. cuspidata* TF sequence was retrieved from the collection of Roche/454 contigs, the reading frame was determined and the authenticity was confirmed by identification of the functional domain. Due to constraints such as the presence of a partial TF domain on truncated contigs or high amino acid divergence from *T. cuspidata* to poplar and *A. thaliana* TFs the possibility cannot be excluded that TF were missed during this analysis. This approach identified a total of 211 TFs within five different TF familys in *T. cuspidata* [Fig.: 4.13 A]. The correlating fold-changes, representative of CMCs, revealed that 115 TFs exhibit d.e. in response to jasmonate in at least one time point of which 60 and 55 show up- and down-regulation respectively [Fig.: 4.13 B] [Appendix V]. Although within the bHLH family 11 and 22 TFs are up- and down-regulated respectively, this family exhibits, with 75 % of all members, the highest jasmonate-responsiveness. 53 % and 51% of the MYB and AP2 TF families respectively show d.e. in response to jasmonate with in both cases significantly more up- than down-regulated TFs. Jasmonate-responsiveness was to a lesser extent observed within the bZIP and WRKY families with predominantly more bZIPs being down-regulated and only 39 % of WRKY TFs in *T. cuspidata* CMCs.

The employed CMCs exhibit a higher production of Taxol^{TM} in response to elicitation by jasmonate [Lee et al., 2010]. Thus gene expression profiles of the subset of up-regulated TFs within CMCs were compared to DDCs. This identified 21 TFs which show higher expression in the employed *T. cuspidata* procambium cell line [Appendix VI]. In an attempt to further characterise the

identified TFs, BLAST searches have been performed to find characterised *A. thaliana* homologues [Appendix VII]. For 10 *T. cuspidata* TFs a homologue could be established within all five TF families. Interestingly five of the characterised TFs are involved in pathogenesis and stress response [homologues of contigs 00499, 22386, 27015, 12425 and 00580] and three in biosynthetic processes [homologues of contigs 17139, 15240 and 27015]. However the homologue of contig 17139 belongs to the AP2 class possessing two AP2 domains. TFs of this subgroup are involved in regulation of developmental processes [Riechmann and Meyerowitz, 1998].

4.6 Discussion

The potent anticancer drug Taxol^{TM} is produced by *T. cuspidata*. Being a not widely used model organism, sequence resources necessary for molecular studies of *T. cuspidata* were not available. Thus sequencing of the transcriptome utilising a combination of Roche/ 454 and Illumina/ Solexa sequencing was used to generate quantitative sequence data and qualitative data, reflecting gene expression profiles respectively under different conditions.

With the aim to detect rare transcripts corresponding to transcriptional regulators, normalisation of the cDNA sample utilised for Roche/ 454 sequencing was performed by Evrogen [Russia]. To eliminate the repetitive sequencing of highly expressed genes, a method involving the denaturation, re-association and degradation using the Kamchatka crab duble-strand nuclease which specifically cleaves nucleic acid duplexes was employed [Zhulidov et al., 2004]. Nor-

malisation of the Roche/ 454 sample resulted in 30,823 unique contigs of a total of 36,905 which corresponds to 17% of repetitive sequencing, which suggests a significant reduction in abundant transcripts. A recent study in *Momordica charantia* found an average increase of gene discovery of 2.9 fold in a Kamchatka crab normalised Roche/ 454 sequenced population when compared to a non-normalised population [Yang et al., 2010].

Gene content in multicellular organisms varies considerably; however, the number of genes in higher plants ranges from 26,500 in *A. thaliana* [Initiative, 2000] to 41,000 and 45,000 in rice and poplar respectively [Sterck et al., 2007]. Assuming a similar gene number in *T. cuspidata* as in the tree poplar with an average transcript length of 1.5 kb [Hilson et al., 2004] a transcriptome coverage using Roche/ 454 technology of 4.4 x was achieved.

One aspect of the transcriptome sequencing concentrated on the identification of MeJA-responsive TF which are expressed frequently at low abundance [Lopato et al., 2006]. Illumina/ Solexa sequencing resulted in the generation of 2 X 10^6 to 16 X 10^6 tags across the utilised samples. This is ample for the detection of low-abundance transcripts which was shown to require a tag count of at least 2 X 10^6 ['t Hoen et al., 2008].

Annotation of the *T. cuspidata* transcriptome resulted in 5,271 contigs being assigned a protein function by BLAST analysis against the plant entries in Uniref100. However 12,902 contigs showed an annotation to an uncharacterised protein and a vast number of 12,650 showed no annotation. The model plant used in this thesis is found within the Taxidae which show a divergence time to other plant models used within the Pinidae of 400 MYA [Sitte et al.,

1991]. Indeed *Picea sitchensis* or Sitka spruce, found in the latter, is the most frequently used plant species annotation of the *T. cuspidata* transcriptome. Although used as a model, *Picea* sp. are less well characterised than the frequently used angiosperm species *A. thaliana*, *O. sativa* and *Z. mays*, which contributed 3.5%, 4.6% and 1.5% respectively to the *T. cuspidata* annotation. This trend was also observed when scoring the sequence identity of distantly related gymnosperm and angiosperm model species.

Confirming the procambium identity

Statistical comparison of the CMC to the DDC transcriptome resulted in the identification of 563 d.e. contigs. This correlates to 1.8 % of the total unique contig number in the *T. cuspidata* transcriptome which are being affected by the characteristics of the employed CMCs. It would be interesting to further analyse these contigs; however, constraints associated with annotation made this initially problematic. Indeed computational annotation was insufficient and after careful manual annotation only 375 of the d.e. contigs were assigned a protein function. These transcripts are likely to contribute to self-renewal and facilitate initiation into the differentiated state of their progeny cells forming the vascular tissue [Singh and Bhalla, 2006]. To this end the employed procambium csc showed the ability to differentiate at high frequency into tracheary elements which are part of the conductive plant tissue under modified culture medium conditions [Lee et al., 2010].

GO term analysis of the d.e. contigs showed that more up-regulated transcripts associated with defence response and response to stress are present

within CMCs. This is consistent with cellular characteristics in mammalian stem cells which exhibit a high resistance to stress [Ramalho-Santos et al., 2002]. This finding can be explained by the need to ensure correct cellular execution as further cell division and subsequent proliferation occurs from stem cell templates [Reya et al., 2001]. Genes associated with lipid metabolism showed a significant down-regulation in CMCs; however, this finding cannot be explained. Although GO term analysis also established a slight up-regulation of contigs associated with signal transduction, Notch or JAK-STAT signalling components involved in regulating stem cell numbers and self-renewal respectively could not be identified among the d.e. contigs [Androutsellis-Theotokis et al., 2006, Kiger et al., 2001].

The homologue of the *A. thaliana PXY* gene shows d.e. in the CMCs. Expression of contig 01805 is significantly up-regulated which suggests a contribution to proper meristem function in *T. cuspidata*. PXY is a receptor-like-kinase, closely related to CLV and BAM123 which confer proliferation and maintenance of dividing cells respectively in the shoot apical meristem [Fisher and Turner, 2007, Clark et al., 1997, DeYoung et al., 2006]. The *pxy* mutant exhibits failure of vascular bundle development and shortened inflouresence in *A. thaliana* indicating its function in vascular tissue formation [Fisher and Turner, 2007].

Maintenance of procambium cells is regulated by auxin and cytokinin signalling in a coordinate fashion [Fukuda, 2004]. Although controlling a variety of developmental processes, brassinosteroids produced in the procambium initiate differentiation of procambial cells in the presence of auxin [Kang et al.,

2010, Clouse and Sasse, 1998]. In this context the *T. cuspidata* homologue, contig 13935, of the *A. thaliana* AIF3 protein is down-regulated in CMCs. AIF3 is a bHLH TF; however, sequence analysis revealed no DNA interaction which suggests a function as a co-regulator. AIF3 was shown to interact with AtBS1 in vitro which in turn stimulates brassinosteroid signalling [Wang et al., 2009]. The domain structure of AIF3 and the highly similar AIF1 suggest a function as negative regulators of AtBS1 [Wang et al., 2009]. This is consistent with contig 13935 being down-regulated in CMCs, which may indicate a role in modulating brassinosteroid mediated differentiation.

Prior to sequencing the *T. cuspidata* transcriptome degenerate PCR and northern analysis was performed to evaluate the expression of the *TED2* homologue from *Z. elegans*. Northern analysis was able to detect a single band in CMCs compared with no signal in DDCs. A previous study showed that during tracheary element differentiation in *Z. elegans* csc specific genes are activated during distinct phases including (I) de-differentiation of somatic cells, (II) the stem-cell-like state and (III) the development of tracheary elements [Fukuda, 1997]. The accumulation of the *TED2* transcript in *Z. elegans* csc during the stem-cell-like-state was restricted to state II and has therefore been utilised as a procambium specific marker gene for *T. cuspidata* CMCs [Fukuda, 1997].

Taken together, the accumulated data suggests that the employed CMCs from *T. cuspidata* conform to the known characteristics of a procambium tissue.

Jasmonate-responsive transcriptome and *T. cuspidata* TFs

The jasmonate elicited subset of samples used was to facilitate the identification of jasmonate-responsive TFs. The expression of Taxol^{TM} structural genes was previously reported in a somatic csc of *T. cuspidata* which peak within 6-12 hours p.e. [Nims et al., 2006]. Thus samples utilised were taken 0.5 h, 2 h and 12 h after jasmonte treatment. Statistical analysis identified 1646 d.e. contigs in at least one of the three time points. This correlates to 5.3 % of the total unique contig number. Dendrogram and heatmap on the clustered 1646 d.e. contigs demonstrate the sample integrity. Although the biological replicates show high overlap in their transcript make-up, CMC#2-2 h p.e. maps out-with the expected position. The underlying cause is unknown. The dendrogram further visualises the genetic reprogramming in *T. cuspidata* csc with the 12 hour samples mapping closest to the untreated cells followed by the 0.5 hour and 2 hour sample post elicitation. This suggests a peak of jasmonate-responsive transcription at 2 hours after treatment.

Annotation limitations didn't allow for further detailed analysis, however differentially regulated transcripts are likely to represent known JA mediated responses including JA signalling components, secondary metabolite structural genes, defence response associated genes and cell cycle components [Pauwels et al., 2008, Galis et al., 2006, Brown et al., 2003]. Further expected are TFs associated with JA-responsive processes. Surprisingly there was only little overlap in jasmonate-responsive gene expression in CMCs and DDCs. A differential effect of jasmonate responses has been previously reported reflecting different cell types of the same species. For instance activation of the

A. thaliana glucosinolate biosynthesis pathway in response to jasmonate was observed in liquid grown seedlings but not in a csc [Sasaki-Sekimoto et al., 2005, Pauwels et al., 2009].

Analysis of the shared d.e. contigs in CMCs and DDCs revealed that the *T. cuspidata* homologues of the jasmonate signalling repressors *JAZ2, JAZ3* and *JAZ10* [reviewed in 1.2.3] show d.e. in response to jasmonate treatment. Whereas *JAZ2* and *JAZ3* are up-regulated in both cell types, *JAZ10* is down-regulated in CMCs. JAZ repressors are important signalling components which liberate the TF MYC2 upon JA perception which in turn activates transcription of early JA response genes including JAZ repressors itself initiating the negative feedback loop [Chini et al., 2007]. The observed fold-change at 2 hours might indicate the peak of MYC2 mediated *JAZ* expression in *T. cuspidata*. Induction of *JAZ* gene expression has been shown in *A. thaliana* in response to jasmonate treatment with the highest expression at 2 hours [Thines et al., 2007]. Interaction of both JAZ2 and JAZ3 with MYC2 has been shown. The latter also interacts with COI1 to mediate JA signalling however JAZ2-COI1 interaction has not yet been determined [Chung et al., 2009]. Expression of *JAZ2* and *JAZ10* has been shown in response to JA and wounding [Yan et al., 2007, Thines et al., 2007] and JAZ10 has been implicated in JA mediated developmental processes, especially in the regulation of cambium during secondary plant growth in *A. thaliana* shoots [Sehr et al., 2010]. How different JAZ proteins and to which extent they mediate specific JA-responsive physiological and metabolic processes is not clear [Chung et al., 2009]. An indication was found by the observation that JAZ-MYC2 interaction depends on a plant-

specific sequence motif which is present only in a small subgroup of bHLH TFs [Chini et al., 2007, Heim et al., 2003]. Thus MYC2 might not be the only JAZ interacting TF [Chung et al., 2009]. It is therefore tempting to speculate a role for JAZ2- and JAZ3-mediated JA induced TF interaction resulting in Taxol^{TM} biosynthesis.

Manual BLAST searches and sequence analysis identified 211 TFs within the AP2/ERF, bHLH, WRKY, MYB and bZIP families. A substantial portion of TFs [115] show d.e. in response to jasmonate, especially the bHLH family. Most frequent up-regulation was observed within the AP2 and MYB families. The AP2 and bHLH TF families operate in jasmonate signalling and their up-regulation, including members of the MYB family has been previously reported [Pauwels et al., 2008]. Expression profile comparison revealed 21 TFs with a higher expression level in response to jasmonate treatment in CMCs. Further homology searches and characterisation of the Tfs established 10 homologues in *A. thaliana* of which are five involved in pathogenesis and stress related processes which might suggest an overlap of JA mediated stress responses. Furthermore two regulators of secondary metabolism could be identified among the *A. thaliana* homologues involved in flavonoid and anthocyanin biosynthesis. Furthermore the 11 remaining TFs for which no *A. thaliana* homologue could be established may function in processes not conserved such as Taxol^{TM} biosynthesis. However these identified TFs [Appendix VII] provide promising targets for functional studies to assess their binding capacity to structural gene promoters thus modulating Taxol^{TM} biosynthesis in CMCs.

Chapter 5

Gene expression profiling of Taxol^{TM} structural genes in CMCs

5.1 Introduction

Biosynthesis of secondary metabolites in plants depends on co-ordinate transcriptional regulation of structural genes [Endt et al., 2002]. This is controlled by specific TFs mediating expression levels of target genes, resulting in regulation of pathway flux [Broun, 2004,Endt et al., 2002]. Indeed, enzymes involved in Taxol^{TM} biosynthesis have been shown to be regulated on the transcriptional level with the two terminal pathway enzymes performing potentially rate-limiting steps [Nims et al., 2006]. These include the attachment of the β-phenylalanine side-chain to baccatin III by BAPT and the addition of a benzoyl group to 3'-*N*-debenzoyl taxol catalysed by DBTNBT [reviewed in 1.5.2; Fig.: 1.4] [Walker et al., 2002, Walker et al., 2002]. To gain further understand-

ing about the transcriptional regulation governing the Taxol™ biosynthetic pathway in *T. cuspidata* and potential TF targets resulting in the observed increased Taxol™ production in CMCs, an insight into the expression profile of structural genes is important.

5.2 *BAPT* is up-regulated in CMCs

As previously mentioned, the employed *T. cuspidata* CMCs exhibit an increased production of Taxol™ [Lee et al., 2010]. RT-PCR was performed to monitor and compare the expression profile of structural genes in CMCs and DDCs. To this end csc of both cell types were adjusted to a culture volume of

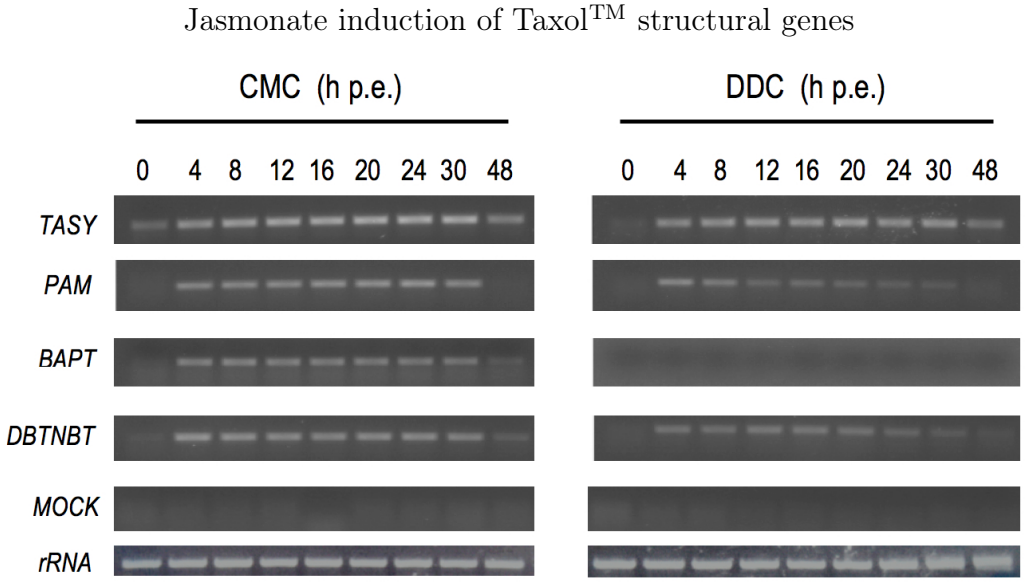

Figure 5.1: Induction of Taxol™ pathway genes in response to jasmonate treatment in CMCs and DDCs over time. Mock: EtOH treatment using PAM primers, rRNA: ribosomal RNA.

50 ml and grown for 2 hours at room temperature in the dark to acclimatise.

The csc were elicited with 100 μM MeJA. Samples were taken over 48 hours with 4 hour increments for up to 24 hours followed by a 30 hour and 48 hour time point. Total RNA was extracted using the RNeasy plant RNA extraction kit [Qiagen, Venlo, Netherlands] followed by oligo-d(T) mediated first strand cDNA synthesis using the Omniscript® RT kit [Qiagen, Venlo, Netherlands] following the manufacturer's instructions.

RT-PCR was performed to assess the expression levels of *TASY*, the first enzyme in the Taxol™ pathway, *BAPT* and *DBTNBT*, which encode the two terminal pathway enzymes and *PAM* [Jennewein et al., 2004]. To aid visualisation a region of interest (ROI) analysis was performed [described in 2.5] which reflects the gene expression level over time.

Expression of *TASY* is highly similar in CMCs and DDCs in the first 20 hours after treatment [Fig.: 5.1 and 5.2 A]. Whereas *TASY* expression peaks at 20 hours in DDCs, expression levels continue to be high in CMCs until 30 hours in response to jasmonate. Expression of *PAM* is up-regulated to a similar level in both cell types at 4 hours p.e. [Fig.: 5.1 and 5.2 B]; however, there is then a steady decline in expression in DDCs. In contrast, expression is maintained to a high level in CMCs until 30 hours after treatment.

The enzyme BAPT, catalysing the second last step en route to Taxol™, shows the highest difference in expression in CMCs to DDCs [Fig.: 5.1 and 5.3 A]. *BAPT* expression is detectable only at a low level 4 - 8 hours after jasmonate treatment in DDCs. In CMCs however, *BAPT* expression peaks at 8 hours and only decreases 30 hours after treatment. There is also a higher expression of *DBTNBT* in CMCs compared to DDCs [Fig.: 5.1 and 5.3 B].

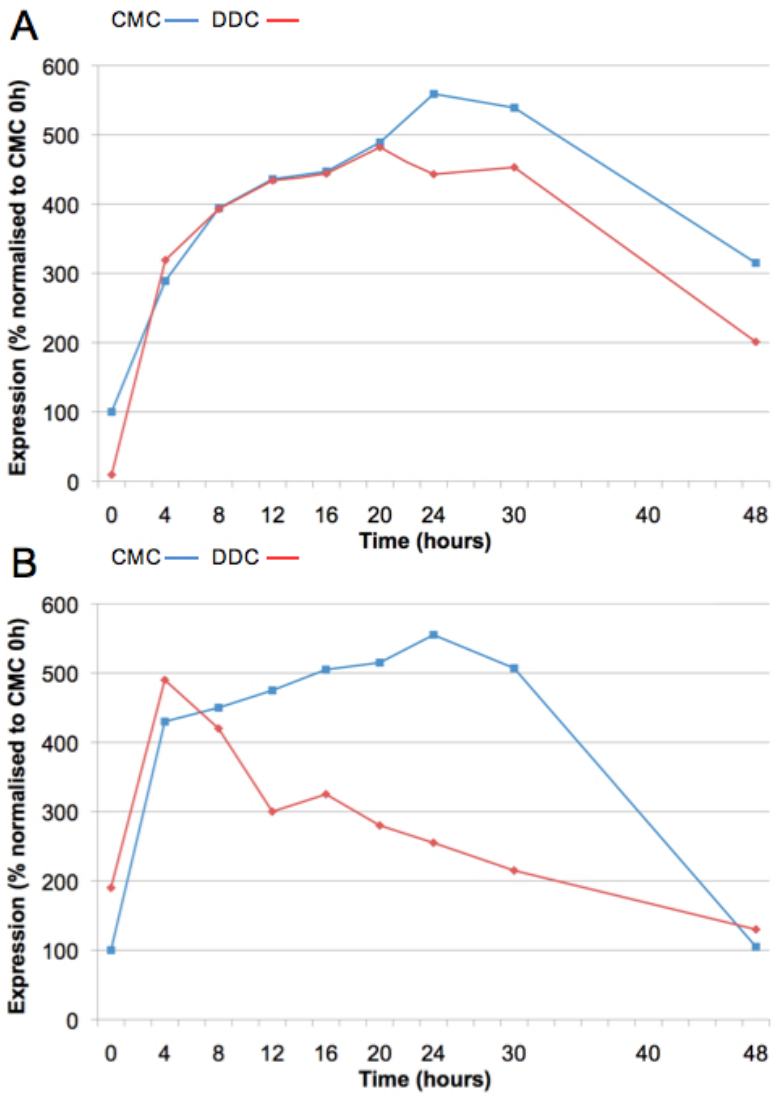

Figure 5.2: Results of the ROI intensity analysis showing the expression of Taxol^{TM} structural genes in response to jasmonate treatment over time. Expression values in % are normalised to the expression in CMC at 0 h p.e. (100%). [A]: *TASY*; [B]: *PAM*.

Both cell types show expression at 4 hours in response to jasmonate although to a higher level in CMCs. *DBTNBT* peaks in DDCs at 12 hours with a constant decrease thereafter. In contrast a steady expression level is observed in CMCs which continues until 30 hours in response to jasmonate treatment.

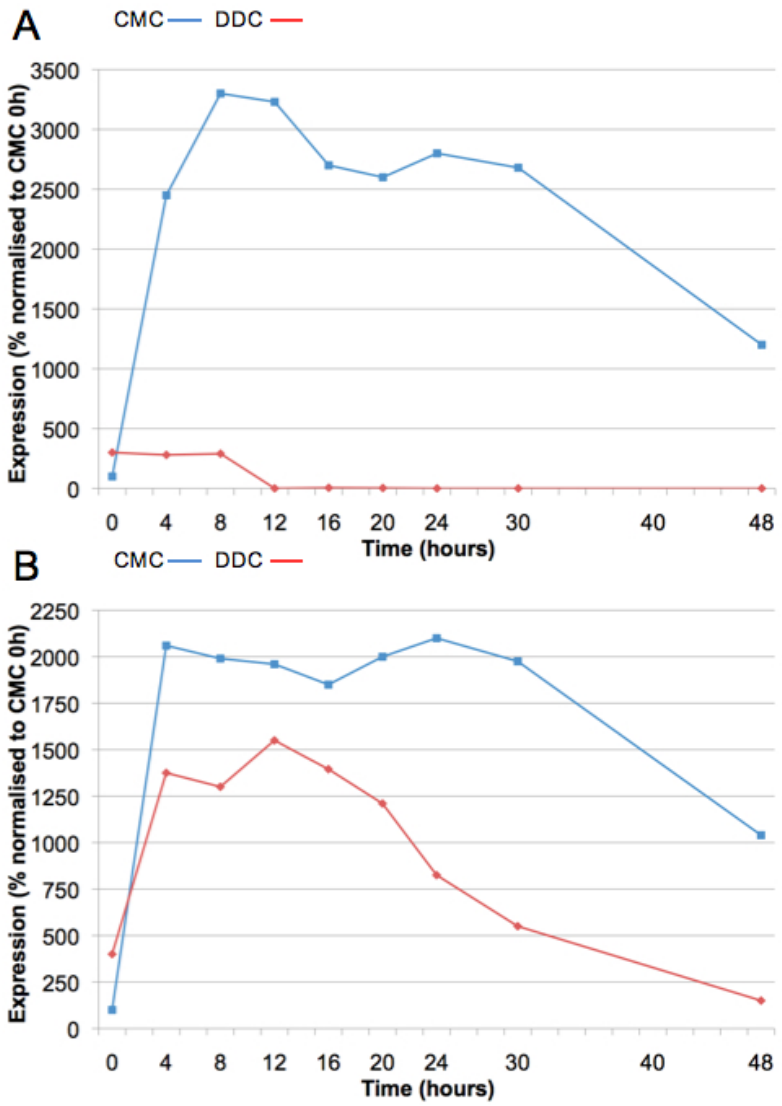

Figure 5.3: Results of the ROI intensity analysis showing the expression of Taxol™ structural genes in response to jasmonate treatment over time. Expression values in % are normalised to the expression in CMC at 0 h p.e. (100%). [A]: *BAPT*; [B]: *DBTNBT*.

5.3 DGE data of Taxol™ structural genes

Taxol™ pathway components were identified among the annotated contigs within the *T. cuspidata* transcriptome [Appendix VIII]. The normalised ex-

pression levels present in CMCs and DDCs of the unelicited and treated samples 12 hours after elicitation were compared [Fig.: 5.4]. Genes involved in the Taxol™ pathway show induction in response to jasmonate, but to different levels of transcript abundance.

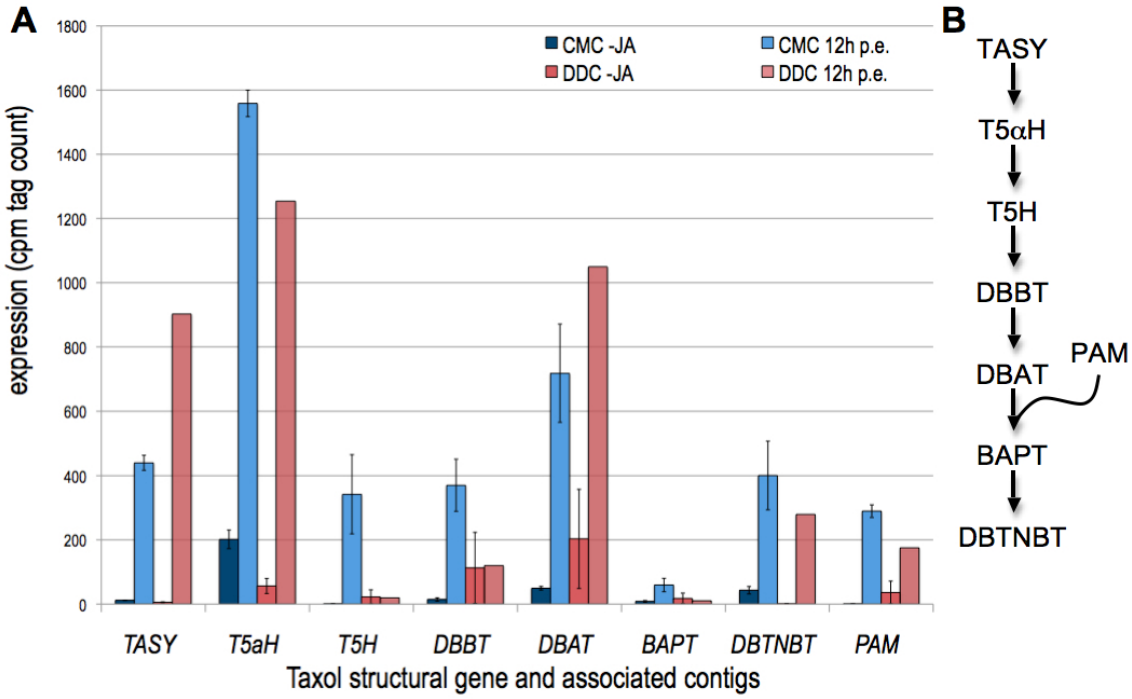

Figure 5.4: [A]: DGE expression profiles of Taxol™ structural genes; [B]: Outline of Taxol™ pathway enzymes.

Expression of the early pathway genes *TASY*, *T5αH* and *T5H* is induced in response to jasmonate 12 hours after treatment; however, transcripts of *T5αH*, especially in CMCs, are present in unelicited cultures. Transcripts of *T5αH* and *T5H* are up-regulated in CMCs. Surprisingly induction of *TASY* is higher in DDCs. In contrast no detectable increase of expression of *T5H* is observed in DDCs.

Although expression levels of *DBBT* and *DBAT* also show up-regulation in response to jasmonate, transcripts of both are observed in unelicited cultures especially in DDCs. Whereas jasmonate treatment induced *DBBT* expression in CMCs, no significant increase of expression is observed in DDCs. Transcript levels of *DBAT* are up-regulated in both cell types in response to jasmonate with a slight increase in abundance in DDCs.

Expression of *BAPT*, *DBTNBT* and *PAM* is up-regulated 12 hours after treatment, however increased expression of *BAPT* is only observed in CMCs. In contrast transcript abundance of *DBTNBT* and *PAM* is increased in both cell types but to a higher level in CMCs.

5.4 Discussion

The employed procambium-derived CMC culture accumulates 443% more Taxol™ in response to jasmonate treatment than a somatic culture of *T. cuspidata* [Lee et al., 2010]. Thus expression levels of the Taxol™ biosynthetic genes in CMCs and DDCs may provide insights in the underlying metabolic regulation. A recent study [Nims et al., 2006] identified *BAPT* and *DBTNBT* as potential bottlenecks within the pathway. Expression level of these two genes along with *TASY* and *PAM* was assessed. *TASY* is the first enzyme in the Taxol™ pathway responsible for the formation of taxa-4(5),11(12)-diene which constitutes the committing step of the pathway [Wildung and Croteau, 1996]. PAM synthesises the Taxol™ side-chain from phenylalanine [Jennewein et al., 2004]. BAPT and DBTNBT catalyse the side-chain attachment to the taxene core

and a benzoylation respectively [Walker et al., 2002, Walker et al., 2002].

RT-PCR revealed an increased transcript abundance in CMCs in response to jasmonate in concert with a differential temporal pattern of expression to DDCs. Transcript abundance of *BAPT* is highly up-regulated after treatment in CMCs. In contrast the same PCR conditions showed only minor amplification in DDCs. The increased expression of *BAPT* may constitute a contributing factor accounting for the observed increased accumulation of Taxol^{TM} in CMCs. Furthermore the expression level of *TASY* in CMCs, although very similar to DDCs, is increased by jasmonate treatment with prolonged transcript accumulation. Similarly expression of *PAM* and *DBTNBT* peaks in DDCs at 4 and 8 hours respectively whereas in CMCs a continued expression is observed until 30 hours after treatment. In conclusion RT-PCR shows that an increased transcript abundance is present within CMCs of all genes monitored, which confirms a regulation of Taxol^{TM} biosynthesis on the transcript level.

DGE expression data from CMCs and DDCs of genes involved in the Taxol^{TM} pathway was compared in the unelicited sample and 12 hours after jasmonate treatment. A similar trend to the RT-PCR data of increased transcript levels in CMCs was observed. However, tag counts of *TASY* show a higher expression in DDCs which could not be confirmed by RT-PCR. Furthermore transcripts levels in the un-elicited cultures for most pathway genes, except *T5αH* are higher in DDCs. It would be interesting to monitor the level of intermediate compounds in both CMCs and DDCs to elucidate whether the observed transcript level correlates to the quantity of intermediates of early

pathway steps in un-elicited cultures. In this context, 10-DAB, the intermediate synthesised by DBBT was found to be present in un-elicited cultures of *T. cuspidata* [Nims et al., 2006], which further suggests that the terminal pathway enzymes confer jasmonate-responsive Taxol^{TM} specificity and the observed basal level of expressed enzymes in DDCs in the absence of jasmonate treatment leads to the synthesis of 10-DAB. In addition levels of Taxol^{TM} accumulation have been found to vary considerably among and within *Taxus* species, depending on factors including plant part and age of plants, and within cultures [Mukherjee et al., 2002, Tabata, 2004]. In this context, a recent report failed to detect Taxol^{TM} by HPLC within a jasmonate elicited culture of *T. chinensis* [Qiu et al., 2009]. The observed irregularity of Taxol^{TM} accumulation could mirror the observed differential levels of transcript abundance detected by RT-PCR and Illumina/ Solexa expression profiling in CMCs and DDCs.

In addition the DGE data reflect the level of transcript abundance at 12 hours after treatment. Although RT-PCR established a continuous transcript level in CMCs from 8 to 30 hours post elicitation, transcript abundance in DDCs may peak earlier as observed with *PAM* and *DBTNBT*. Thus the 12 hour time point does not constitute an actual representation of transcript abundance in DDCs over time.

Chapter 6

General discussion

The secondary metabolite and potent anti-cancer drug Taxol^{TM} is synthesised by all species of *Taxus* (yew) [Croteau et al., 2006]. Of hugh pharmaceutical importance, Taxol^{TM} has been shown to be effective against a variety of cancers. Furthermore the FDA approved the plant cell culture process for supply of Taxol^{TM}, moving away from the unsustainable extraction of Taxol^{TM} from bark and semisynthesis [Nims et al., 2006, Goodman and Walsh, 2001].

To this end a csc from *T. cuspidata* was employed as the model organism utilised in this thesis. Initiated from isolated procambium cells the cell line consists of homogenous undifferentiated CMCs [Lee et al., 2010]. Favourable characteristics such as an increased production of Taxol^{TM}, fast and stable growth rate and limited cell aggregation in comparison to a somatic DDC line from *T. cuspidata* [Lee et al., 2010] make CMCs an optimal model to study the transcriptional regulation of the Taxol^{TM} biosynthetic pathway. In this work standard molecular approaches were adapted to *T. cuspidata* and high

throughput sequencing Roche/ 454 and Illumina/ Solexa have been applied to confirm the procambium characteristics of CMCs and to identify putative transcriptional activators governing Taxol™ biosynthesis.

Taxol™ pathway regulation

The plant hormone jasmonate mediates stress induced synthesis of secondary metabolites in various plant species [Memelink et al., 2001]. More importantly transcripts encoding for enzymes involved in the Taxol™ pathway show increased expression in response to methyljasmonate treatment resulting in accumulation of Taxol™ in *Taxus* csc [Mirjalili and Linden, 1996, Nims et al., 2006]. In order to identify jasmonate signalling components and transcriptional activators governing the Taxol™ pathway the transcriptome sequencing of *T. cuspidata* was performed employing Roche/ 454 in combination with Illumina/ Solexa sequencing. Comparison of d.e. contigs within CMCs and DDCs uncovered differential expression of the *T. cuspidata* homologues of the jasmonate signalling components JAZ2 and JAZ3 in response to jasmonate treatment. Consistent with MYC2 mediated *JAZ* expression in *A. thaliana* [Thines et al., 2007], the highest fold-change in *T. cuspidata* was observed at 2 hours following treatment. JAZ proteins act as suppressors of the MYC2 TF which activates the expression of early jasmonate-responsive genes including JAZ suppressors [Chini et al., 2007]. Interaction of both JAZ2 and JAZ3 with MYC2 was shown; however, Chung et al., [2009] argued that on the basis of JAZ-MYC2 interaction characteristics, other TFs may also be JAZ protein targets responsible for mediating the diverse jasmonate-responsive processes.

Thus the JA-signalling components JAZ2 and JAZ3 might be involved in mediating transcriptional regulation of Taxol™ biosynthesis.

Blast searches and sequence analysis identified 211 TFs within the AP2/ERF, bHLH, bZIP, WRKY and MYB families of which 115 show differential expression in response to jasmonate. TFs within the AP2/ERF family showed

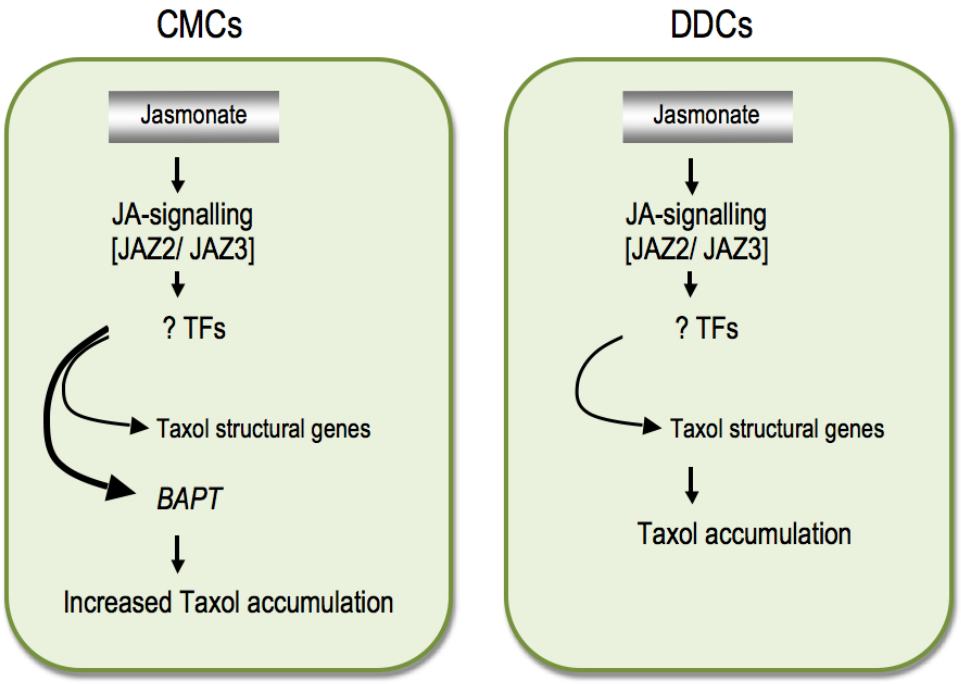

Figure 6.1: Model for regulation of the Taxol™ pathway in CMCs: jasmonate treatment in CMCs and DDCs triggers the synthesis of Taxol™ structural genes mediated by JA-signalling including the MYC2 suppressors JAZ2 and JAZ3 through the transcriptional activation of unknown TFs. A higher expression level of *BAPT* may contributes to a increased Taxol™ accumulation in CMCs.

the highest up-regulation in response to jasmonate treatment. Members of the AP2 family such as the ORCA3 TF in *C. roseus*, which mediates transcriptional activation of TIA biosynthesis, are jasmonate inducible [van der

Fits and Memelink, 2001]. As the CMC suspension culture exhibits an increased amount of Taxol^{TM} accumulation, expression levels of the subset of up-regulated TFs in response to treatment in CMCs were compared with DDCs. This revealed 21 TF candidates which provide promising targets for further functional analysis. The largest number of d.e. TFs which show higher expression in CMCs belong to the AP2/ERF [10 TFs] and the MYB [5 TFs] families. Both have a well characterised function in the transcriptional regulation of secondary metabolism [Menke et al., 1999, Galis et al., 2006]. MYB TFs are known to hetero-dimerise with TFs of the bHLH family [Mol et al., 1998], two members of which show also higher expression in CMCs. Homology searches established 11 *A. thaliana* homologues of which two are involved in the regulation of secondary metabolism. Five contigs were found to function in pathogenesis or response to stress which might be consistent with observed JA induced stress-responsiveness [Reymond and Farmer, 1998]. Furthermore, TFs are valuable tools for the engineering of metabolic pathways. For example this was highlighted by over-expression of the TFs C1 and LEAF COLOUR in the crop plant tomato which resulted in an increased accumulation of health beneficial flavonols [Bovy et al., 2002]. Once interaction of *T. cuspidata* TFs with promoters of Taxol^{TM} structural genes has been confirmed, over-expression of the regulators may result in an increased Taxol^{TM} biosynthesis. Further analysis of the translated TF amino acid sequence of identified TFs was impeded by a large number of truncated transcripts. Phylogenetic analysis could reveal subgroups within families and uncover evolutionarily conserved or *Taxus* specific domain characteristics.

After attempts to construct a BAC library from *T. cuspidata* genomic DNA, genome walking led to the isolation of a 513 bp long fragment corresponding to the promoter region of *TASY*. Subsequent in-silico analysis of the *TASY* and *DBTNBT* promoter established the presence of multiple jasmonate-responsive [GCC-box, G-box and T/G-box] and pathogen inducible [W-box] promoter elements. This is consistent with Taxol^{TM} accumulation in response to jasmonate and fungal elicitation [Mirjalili and Linden, 1996, Xu et al., 2004]. Further jasmonate inducibility could be shown for the *TASY* promoter in *N. benthamiana*. However *DBTNBT* expression in response to jasmonate treatment could not be detected. This may indicate that factors are required which are not evolutionarily conserved from *N. benthamiana* to *Taxus*. Dual infiltration of identified regulators and chimeric promoter-reporter constructs in the presence and absence of jasmonate would further establish an interaction of TFs to Taxol^{TM} structural genes.

RT-PCR revealed that transcripts of *TASY, PAM, BAPT* and *DBTNBT* accumulate in CMCs and DDCs in response to jasmonate treatment. The higher and prolonged expression level in CMCs might be indicative of increased Taxol^{TM} accumulation. A further contributing factor to the latter could be the increased expression of *BAPT* in CMCs versus DDCs. The reaction performed by BAPT was previously identified as a potential rate limiting step, ligating the phenylalanine derived side-chain to the taxene core [Nims et al., 2006, Walker et al., 2002].

Procambium associated genes are expressed in CMCs

The procambium is a primary meristem involved in the formation of vascular tissue [Fukuda, 2004]. Vascular development from procambium cells involves the differentiation of phloem and xylem precursor cells into the distinct cell types forming the vascular bundles. Present among them are TEs which undergo cell death to form hollow tubes for movement of fluids [Fukuda, 2004].

Crucial in the development of vascular bundles is ordered cell division. The receptor like kinase PXY is expressed within dividing meristematic procambium cells of *A. thaliana* where it is essential for ordered cell division [Fisher and Turner, 2007]. Consistent with this is the up-regulation of the *PXY* homologue in *T. cuspidata* CMCs. Receptor like kinases have also been implicated in the asymmetric cell division prior to lateral root initiation in *A. thaliana* [Smet et al., 2008]. However in the ordered cell division of vascular initials PXY interacts with the peptide ligands CLE41 and CLE42 which are expressed in adjoining phloem cells thus performing orientation dependent signalling resulting in co-ordinated cell division [Etchells and Turner, 2010].

Mesophyll cells of *Z. elegans* can differentiate into tracheary elements in culture. This transdifferentiation (dedifferentiation followed by redifferentiation) process involves three distinct stages [Fig.: 5.1] in which mesophyll cells dedifferentiate (I) followed by a procambium-like state (II) which corresponds to potency restriction and transition into precursor cells of tracheary elements and (III) the formation of a secondary wall and developmentally programmed cell death [Fukuda, 1997]. Transcripts which accumulate in the procambium-like stage II include *TED2* which encodes a ζ-crystallin hydrophobic polypep-

tide [Fukuda, 1997, Demura and Fukuda, 1994]. Northern analysis showed *TED2* expression in *T. cuspidata* CMCs. In contrast no expression was detected in DDCs. Expression analysis in *Z. elegans* seedlings showed that the

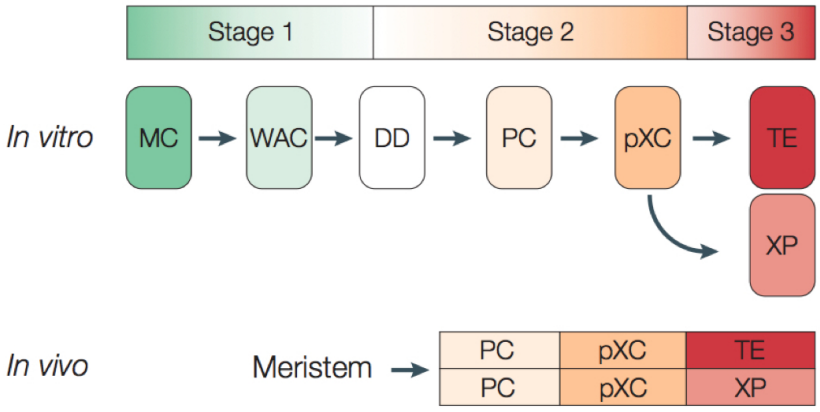

Figure 6.2: Transdifferentiation stages of *Z. elegans* somatic cells to tracheary elements. MC: mesophyll cells, WAC: wound-activated cells, DD: dedifferentiated cells, PC: procambial cells, pXC: xylem cell precursors, TE: tracheary elements, XP: xylem parenchyma cells [Fukuda, 2004].

TED2 transcript accumulates in a very early phase in procambial tissue in immature primary xylem and phloem cells within the root/ hypocotyl boundary and in root procambium cells [Demura and Fukuda, 1994]. This indicates that the *T. cuspidata TED2* transcript detected in CMCs is indicative of the procambium characteristic of this csc.

Maintenance and differentiation of procambium cells within the plant tissue context is regulated by polar auxin transport, cytokinin and brassenosteroids signalling [Fukuda, 2004]. Cytokinin was implicated in the formation of procambium by mutation in the *A. thaliana wooden leg* locus. Mutants exhibit a reduced number of procambium cells in embryos [Scheres et al., 1995].

WOODEN LEG was subsequently characterised as a cytokinin receptor [Maehoenen et al., 2000]. A *T. cuspidata* homologue of *WOODEN LEG* was found to be up-regulated 12 fold in CMCs relative to DDCs [Lee et al., 2010] which suggests a role for cytokinin in the maintenance of *T. cuspidata* procambium cells.

Intracellular signals which confer procambium maintenance and initiate differentiation into xylem precursor cells include auxin and BR signalling [Fukuda, 2004]. Furthermore the effect of BR signalling on xylem formation was demonstrated in cress seedlings which showed reduced xylem but increased phloem differentiation in response to a BR biosynthesis inhibitor [Nagata et al., 2001]. Endogenous BR content increases in transdifferentiating *Z. elegans* cells which coincides with the transition from procambium-like cells to tracheary elements (stage 2 to stage 3 in vitro) [Fig.: 6.1] [Yamamoto et al., 2001]. The BR signal triggers transcriptional activation of xylem-precursor cell related gene expression [Fukuda, 2004]. AtBS1, a bHLH protein was recently shown to function as a stimulator of BR signalling and AIF3 a negative regulator of AtBS1 was identified in an yeast two-hybrid screen [Wang et al., 2009]. The *T. cuspidata* homologue of AIF3 is significantly down-regulated in CMCs which would suggest increased BR signalling in CMCs indicating an overlap with transdifferentiating *Z. elegans* stage II cells. Thus contig 13935 might be involved in mediating BR signalling in *T. cuspidata* procambium cells. Investigations into how AIF proteins function are required [Wang et al., 2009]. Furthermore key components of BR signalling might be conserved from *A. thaliana* and *Z. elegans* to *T. cuspidata*, however, the presence of additional

factors modulating BR mediated maintenance and differentiation of CMCs cannot be ruled out. It would be interesting to measure the BR level in both cell types to elucidate whether the down-regulation of AIF3 homologue in *T. cuspidata* CMCs mirrors the BR content.

CMCs exhibit increased defences

Mapping of the most frequent GO terms among the 563 d.e. contigs in CMCs versus DDCs revealed that the subgroup of contigs associated with "response to stress" and "defence response" are up-regulated in CMCs. Consistent with this finding is the high resistance to stress which was reported in mammalian stem cells [Ramalho-Santos et al., 2002]. Furthermore, in plants two theories have been proposed regarding the accumulation of protective secondary metabolites such as the diterpenoid Taxol^{TM}. Tissues with the highest fitness value such as meristems are predicted, according to the "optimal defence theory", to receive the highest level of protection [Rhoades, 1979]. In contrast the "growth differentiation balance hypothesis" predicts an opposite pattern for the presence of protective compounds, as growth (performed by sink-tissues) precedes the differentiation process (performed by source-tissues) [Herms and Mattson, 1992,Cronin and Hay, 1996]. Content of the antiherbivore phlorotannin compounds in kelps and rockweed was found to be higher in meristematic tissue compared to vegetative tissue which supports the "optimal defence theory" [VanAlstyne et al., 1999]. Consistent with this is a higher level of phenolic compounds of 8 mg/ g dry weight in CMCs versus 0.7 mg/ g dry weight in DDCs [Lee, personal communication]. Phenolics and terpenoids are the most

important defence compounds in conifers [Mumm and Hilker, 2006]. The observed 443% increase in Taxol™ synthesis in CMCs might also be attributed to an increased defence strategy of *T. cuspidata* meristematic cells. Furthermore the increased amounts of phenolics and Taxol™ are present within a homogenous csc indicating a translocation-independent system contrasting allocation within the plant tissue context. This suggests that innate cellular characteristics result in increased biosynthesis of protective compounds.

Considerations for improved Taxol™ synthesis in CMCs

Plant cell cultures from *Taxus* species are an attractive source for Taxol™ supply. Advantages are a faster growth rate and a lower content of waxy constituents, pigments and non-polar lipids compared to the bark or plant material of *Taxus*, which simplifies the extraction of Taxol™ [Jaziri et al., 1996]. The culture environment and medium composition can be tightly controlled resulting in optimal growth conditions [Leistner, 2005]. The utilised CMCs in this thesis exhibit an increased production of the secondary metabolite [Lee et al., 2010]. A further possible explanation for this can be found when considering previously conducted Taxol™ localisation experiments. Using an antitaxol antiserum, Taxol™ within tissues of *T. cuspidata* was shown to be almost exclusively located in the cell walls of phloem, vascular cambium and xylem [Russin et al., 1995]. Although Taxol™ accumulation is reported in csc of both CMCs and DDCs, [Lee et al., 2010], procambium cells give rise to vascular tissues where Taxol™ was predominantly localised, it is therefore tempting to speculate that factors are present within procambium cells

which are a prerequisite for a high Taxol^{TM} content within the *Taxus* tissue context. These could have been retained within CMCs resulting in increased production of Taxol^{TM}. Attempts have been made in this thesis to identify contributing up-regulated transcriptional regulators and rate limiting steps within the Taxol^{TM} pathway in CMCs. In conclusion CMCs constitute an ideal system for the supply of Taxol^{TM}. This however could be further improved by the isolation of CMCs from *T. floridana* which has a higher Taxol^{TM} content of 516 μg/ g dry needles compared to *T. cuspidata* with only 105 μg/ g dry needles [van Rozendaal et al., 2000].

For engineering of metabolic pathways, stable plant transformation techniques are necessary. *Agrobacterium* mediated transformation of shoot segments and csc was shown with varying success in several *Taxus* species including *T. cuspidata* [Han et al., 1994, Ketchum et al., 2007]. To increase pathway flux the over expression of rate-limiting enzymes has been successfully applied to improve the accumulation of essential oils in *Mentha x piperita L.* [Mahmoud and Croteau, 2001]. Especially BAPT was previously identified as performing a rate-limiting step [Nims et al., 2006]. DGE expression data indicates a low expression of *BAPT* even in jasmonate treated CMCs compared to other pathway genes, which makes this gene a target in an over-expression approach. However, increased accumulation of a desired compound may also depend on appropriate precursor and co-factor supply, storage or translocation systems [Zulak et al., 2007]. These aspects, including the regulation of structural genes, are controlled by TFs [Endt et al., 2002]. To this end 211 TFs within five families have been identified and partially characterised, which con-

stitutes candidates for further functional elucidation and eventually to increase Taxol™ pathway flux.

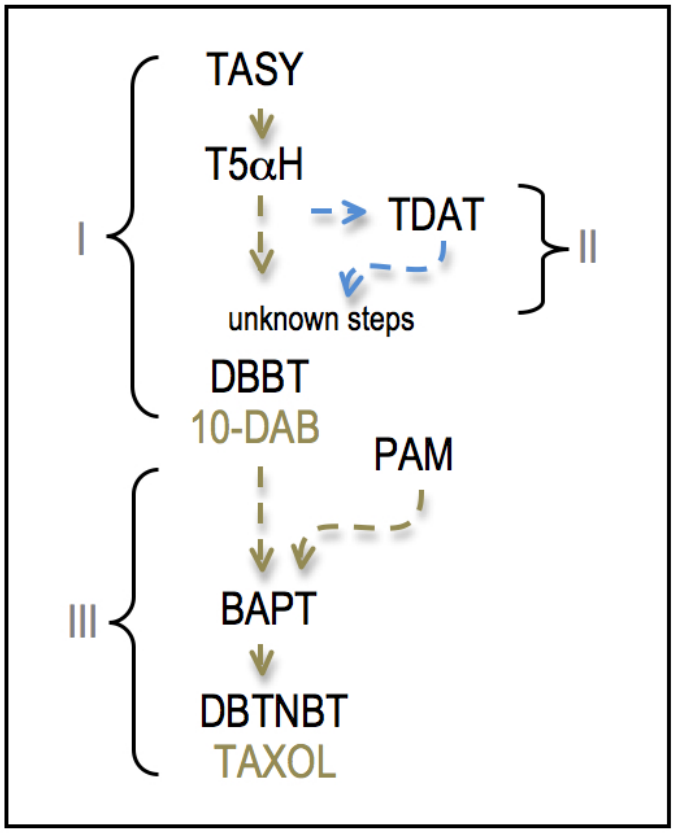

Figure 6.3: Simplified version of Taxol™ pathway branches. The pathway follows the route highlighted in green for the biosynthesis of Taxol™. I: partially jasmonate independent steps, II: steps not involved in Taxol™ synthesis, III: late pathway steps. Enzymes (black), Intermediates (green and blue). Adapted from [Nims et al., 2006].

The Taxol™ pathway can be divided into several segments [Nims et al., 2006]. *TASY* to *DBBT* represent (I) early pathway steps [Fig.: 6.3], (II) a branch point implementing *TDAT* and (III) the late pathway steps from the intermediate 10-DAB to Taxol™ which also includes synthesis and ligation of the side chain. 10-DAB was detected in unelicited cultures [Nims et al.,

2006] indicating that the early pathway is partially jasmonate independent. The branch point enzymes are not involved in the synthesis of Taxol™ [Nims et al., 2006]. This suggests that several modes of transcriptional regulation (which may share common factors) must be in place which govern the distinct pathway segments. Similarly differential expression of two early pathway genes involved in anthocyanin biosynthesis in *Z. mays* in response to over-expression of the TFs *R* and *C1* was observed [Grotewold et al., 1998]. This indicates that for complex pathway regulation more than one TF may be required.

A further consideration in improving Taxol™ biosynthesis could involve *T. cuspidata* homologues of known jasmonate-responsive TFs from other species. For this, however, evolutionary conservation/ divergence from angiosperm species to *T. cuspidata* must be factored into investigation. Attempts to identify homologues of the *C. roseus* ORCA TFs in *T. cuspidata* were not successful. ORCA3 governs jasmonate-responsive transcriptional regulation of genes in primary and secondary metabolism involved in TIA biosynthesis [van der Fits and Memelink, 2000]. There is a divergence time of 400 MY in which angiosperms to *Taxus* species co-evolved which might explain the absence of ORCA homologues. Furthermore secondary metabolism is diverse and adaptive [Hartman, 1996] thus genes and regulators involved are under less conservation pressure. A diverged evolution resulting in species specific regulation was shown in maize and petunia on genes involved in anthocyanin synthesis. In response to TF over-expression, the *C2* gene in maize and *CHALCONE SYNTHASE A* in petunia are induced and not affected respectively, which indicates a modification within their promoter sequences [Quattrocchio

et al., 1998]. Therefore even if counterparts of known JA-responsive TFs involved in secondary metabolism from angiosperm species within *T. cuspidata* exist, the regulators or their targets may have functionally diverged.

References

[IUC, 2010] (2010). The IUCN Red List of Threatened Species.

[Agrawal et al., 2002] Agrawal, A. A., Conner, J. K., Johnson, M. T. J. and Wallsgrove, R. (2002). Ecological genetics of an induced plant defense against herbivores: additive genetic variance and costs of phenotypic plasticity. Evolution *56*, 2206–2213.

[Allina et al., 1998] Allina, S. M., Pri-Hadash, A., Theilmann, D. A., Ellis, B. E. and Douglas, C. J. (1998). 4-Coumarate:coenzyme A ligase in hybrid poplar. Properties of native enzymes, cDNA cloning, and analysis of recombinant enzymes. Plant Physiol *116*, 743–754.

[Amoutzias et al., 2007] Amoutzias, G. D., Veron, A. S., Weiner, J., Robinson-Rechavi, M., Bornberg-Bauer, E., Oliver, S. G. and Robertson, D. L. (2007). One billion years of bZIP transcription factor evolution: conservation and change in dimerization and DNA-binding site specificity. Mol Biol Evol *24*, 827–835.

[Androutsellis-Theotokis et al., 2006] Androutsellis-Theotokis, A., Leker, R. R., Soldner, F., Hoeppner, D. J., Ravin, R., Poser, S. W., Rueger, M. A., Bae, S., Kittappa, R. and McKay, R. D. G. (2006). Notch signalling regulates stem cell numbers in vitro and in vivo. Nature *442*, 823–826.

[Arnal and Wade, 1995] Arnal, I. and Wade, R. H. (1995). How does taxol stabilize microtubules? Curr Biol *5*, 900–908.

[Atchley and Fitch, 1997] Atchley, W. R. and Fitch, W. M. (1997). A natural classification of the basic helix-loop-helix class of transcription factors. Proc Natl Acad Sci U S A *94*, 5172–5176.

[Baker et al., 2010] Baker, C. M., Chitrakar, R., Obulareddy, N., Panchal, S., Williams, P. and Melotto, M. (2010). Molecular battles between plant and pathogenic bacteria in the phyllosphere. Braz J Med Biol Res *43*, 698–704.

[Baldwin and Osheroff, 2005] Baldwin, E. L. and Osheroff, N. (2005). Etoposide, topoisomerase II and cancer. Curr Med Chem Anticancer Agents *5*, 363–372.

[Baldwin, 1998] Baldwin, I. T. (1998). Jasmonate-induced responses are costly but benefit plants under attack in native populations. Proc Natl Acad Sci U S A *95*, 8113–8118.

[Baloglu and Kingston, 1999] Baloglu, E. and Kingston, D. G. I. (1999). The taxane diterpenoids. J Nat Prod *62*, 1448–1472.

[Balunas and Kinghorn, 2005] Balunas, M. J. and Kinghorn, A. D. (2005). Drug discovery from medicinal plants. Life Sci *78*, 431–441.

[Bates, 1999] Bates, G. W. (1999). Plant transformation via protoplast electroporation. Methods Mol Biol *111*, 359–366.

[Benedetti et al., 1995] Benedetti, C. E., Xie, D. and Turner, J. G. (1995). Coi1-dependent expression of an *Arabidopsis* vegetative storage protein in

flowers and siliques and in response to coronatine or methyl jasmonate. Plant Physiol *109*, 567–572.

[Bennett et al., 1994] Bennett, R. N., Wallsgrove, R. and Wallsgrove, M. (1994). Secondary metabolites in plant defence mechanisms. New Phytologist *127*, 617–633.

[Blagosklonny, 2002] Blagosklonny, M. V. (2002). P53: an ubiquitous target of anticancer drugs. Int J Cancer *98*, 161–166.

[Blagosklonny and Fojo, 1999] Blagosklonny, M. V. and Fojo, T. (1999). Molecular effects of paclitaxel: myths and reality (a critical review). Int J Cancer *83*, 151–156.

[Blanchette, 1992] Blanchette, R. A. (1992). Anatomical responses of xylem to injury and invasion by fungi. In Defense mechanisms of woody plants against fungi. Springer Verlag, Berlin.

[Bolta et al., 2003] Bolta, Z., Baricevic, D. and Raspor, P. (2003). Biomass segregation in sage cell suspension culture. Biotechnol Lett *25*, 61–65.

[Boter et al., 2004] Boter, M., Ruiz-Rivero, O., Abdeen, A. and Prat, S. (2004). Conserved MYC transcription factors play a key role in jasmonate signaling both in tomato and *Arabidopsis*. Genes Dev *18*, 1577–1591.

[Bouvier et al., 2005] Bouvier, F., Rahier, A. and Camara, B. (2005). Biogenesis, molecular regulation and function of plant isoprenoids. Prog Lipid Res *44*, 357–429.

[Bovy et al., 2002] Bovy, A., de Vos, R., Kemper, M., Schijlen, E., Pertejo, M. A., Muir, S., Collins, G., Robinson, S., Verhoeyen, M., Hughes, S., Santos-Buelga, C. and van Tunen, A. (2002). High-flavonol tomatoes resulting from the heterologous expression of the maize transcription factor genes LC and C1. Plant Cell *14*, 2509–2526.

[Briggs et al., 2007] Briggs, A. W., Stenzel, U., Johnson, P. L. F., Green, R. E., Kelso, J., Pruefer, K., Meyer, M., Krause, J., Ronan, M. T., Lachmann, M. and Paeaebo, S. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. Proc Natl Acad Sci U S A *104*, 14616–14621.

[Broun, 2004] Broun, P. (2004). Transcription factors as tools for metabolic engineering in plants. Curr Opin Plant Biol *7*, 202–209.

[Brown et al., 2003] Brown, R. L., Kazan, K., McGrath, K. C., Maclean, D. J. and Manners, J. M. (2003). A role for the GCC-box in jasmonate-mediated activation of the PDF1.2 gene of *Arabidopsis*. Plant Physiol *132*, 1020–1032.

[Buchholz et al., 1983] Buchholz, K., Rapp, P. and Zadrazil, F. (1983). Methods of Enzymatic Analysis Vol. II, 178-80,. Verlag Chemie, Weinheim.

[Buck and Atchley, 2003] Buck, M. J. and Atchley, W. R. (2003). Phylogenetic analysis of plant basic helix-loop-helix proteins. J Mol Evol *56*, 742–750.

[Butler, 2005] Butler, M. S. (2005). Natural products to drugs: natural product derived compounds in clinical trials. Nat Prod Rep *22*, 162–195.

[Cai et al., 2008] Cai, D. H., Wang, D., Keefer, J., Yeamans, C., Hensley, K. and Friedman, A. D. (2008). C/EBP alpha:AP-1 leucine zipper heterodimers

bind novel DNA elements, activate the PU.1 promoter and direct monocyte lineage commitment more potently than C/EBP alpha homodimers or AP-1. Oncogene *27*, 2772–2779.

[Camehl and Oelmueller, 2010] Camehl, I. and Oelmueller, R. (2010). Do ethylene response factors-9 and -14 repress PR gene expression in the interaction between *Piriformospora indica* and *Arabidopsis*? Plant Signal Behav *5*.

[Cheong et al., 2003] Cheong, Y. H., Moon, B. C., Kim, J. K., Kim, C. Y., Kim, M. C., Kim, I. H., Park, C. Y., Kim, J. C., Park, B. O., Koo, S. C., Yoon, H. W., Chung, W. S., Lim, C. O., Lee, S. Y. and Cho, M. J. (2003). BWMK1, a rice mitogen-activated protein kinase, locates in the nucleus and mediates pathogenesis-related gene expression by activation of a transcription factor. Plant Physiol *132*, 1961–1972.

[Chini et al., 2007] Chini, A., Fonseca, S., FernÃąndez, G., Adie, B., Chico, J. M., Lorenzo, O., Garcia-Casado, G., Lopez-Vidriero, I., Lozano, F. M., Ponce, M. R., Micol, J. L. and Solano, R. (2007). The JAZ family of repressors is the missing link in jasmonate signalling. Nature *448*, 666–671.

[Cho et al., 2008] Cho, K., Wang, X., Nie, S., Chen, Z. G. and Shin, D. M. (2008). Therapeutic nanoparticles for drug delivery in cancer. Clin Cancer Res *14*, 1310–1316.

[Chung et al., 2009] Chung, H. S., Niu, Y., Browse, J. and Howe, G. A. (2009). Top hits in contemporary JAZ: an update on jasmonate signaling. Phytochemistry *70*, 1547–1559.

[Clark et al., 1997] Clark, S. E., Williams, R. W. and Meyerowitz, E. M. (1997). The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in *Arabidopsis*. Cell *89*, 575–585.

[Clouse and Sasse, 1998] Clouse, S. D. and Sasse, J. M. (1998). BRASSINOSTEROIDS: Essential regulators of plant growth and development. Annu Rev Plant Physiol Plant Mol Biol *49*, 427–451.

[Cocking, 1972] Cocking, E. C. (1972). Plant Cell Protoplasts - Isolation and Development. Ann. Rev. Plant Physiol. *23*, 29–50.

[Coemans et al., 2005] Coemans, B., Matsumura, H., Terauchi, R., Remy, S., Swennen, R. and Sagi, L. (2005). SuperSAGE combined with PCR walking allows global gene expression profiling of banana (*Musa acuminata*), a non-model organism. Theor Appl Genet *111*, 1118–1126.

[Connolly and Hill, 1991] Connolly, J. D. and Hill, R. A. (1991). Dictionary of Terpenoids. Chapman and Hall, London.

[Corbineau et al., 1988] Corbineau, F., Rudnicki, R. M. and Come, D. (1988). The effects of methyl jasmonate on sunflower (*Helianthus annuus L.*) seed germination and seedling development. Plant Growth Regulation *7*, 157–169.

[Cragg, 1998] Cragg, G. M. (1998). Paclitaxel (Taxol): a success story with valuable lessons for natural product drug discovery and development. Med Res Rev *18*, 315–331.

[Cronin and Hay, 1996] Cronin, G. and Hay, M. E. (1996). Within-plant variation in seaweed palatability and chemical defenses: optimal defense theory versus the growth-differentiation balancehypothesis. Oecologia *105*, 361–368.

[Croteau et al., 2006] Croteau, R., Ketchum, R. E. B., Long, R. M., Kaspera, R. and Wildung, M. R. (2006). Taxol biosynthesis and molecular genetics. Phytochem Rev *5*, 75–97.

[Crown and O'Leary, 2000] Crown, J. and O'Leary, M. (2000). The taxanes: an update. Lancet *355*, 1176–1178.

[Curtis and Grossniklaus, 2003] Curtis, M. D. and Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol *133*, 462–469.

[Dathe et al., 1981] Dathe, W., Roensch, H., Preiss, A., Schade, W., Sembdner, G. and Schreiber, K. (1981). Endogenous plant hormones of the broad bean, *Vicia faba L.* (-)-jasmonic acid, a plant growth inhibitor in pericarp. PLANTA *153*, 530–535.

[Davey et al., 2005] Davey, M. R., Anthony, P., Power, J. B. and Lowe, K. C. (2005). Plant protoplasts: status and biotechnological perspectives. Biotechnol Adv *23*, 131–171.

[Davies and Schwinn, 2003] Davies, K. M. and Schwinn, K. E. (2003). Transcriptional regulation of secondary metabolism. Functional Plant Biology *30*, 913–925.

[de Laubenfels, 1988] de Laubenfels, D. J. (1988). Flora Malesiana: Coniferales. Dordrecht: Kluwer Academic.

[del Pozo and Estelle, 2000] del Pozo, J. C. and Estelle, M. (2000). F-box proteins and protein degradation: an emerging theme in cellular regulation. Plant Mol Biol *44*, 123–128.

[Demura and Fukuda, 1994] Demura, T. and Fukuda, H. (1994). Novel vascular cell-specific genes whose expression is regulated temporally and spatially during vascular system development. Plant Cell *6*, 967–981.

[Denis et al., 1988] Denis, J. N., Greene, A. E., Guenard, D., Gueritte-Voegelein, F., Mangatal, L. and Potier, P. (1988). Highly efficient, practical approach to natural taxol. J. Am. Chem. Soc *110*, 5917.

[Deppmann et al., 2006] Deppmann, C. D., Alvania, R. S. and Taparowsky, E. J. (2006). Cross-species annotation of basic leucine zipper factor interactions: Insight into the evolution of closed interaction networks. Mol Biol Evol *23*, 1480–1492.

[Devoto et al., 2002] Devoto, A., Nieto-Rostro, M., Xie, D., Ellis, C., Harmston, R., Patrick, E., Davis, J., Sherratt, L., Coleman, M. and Turner, J. G. (2002). COI1 links jasmonate signalling and fertility to the SCF ubiquitin-ligase complex in *Arabidopsis*. Plant J *32*, 457–466.

[DeYoung et al., 2006] DeYoung, B. J., Bickle, K. L., Schrage, K. J., Muskett, P., Patel, K. and Clark, S. E. (2006). The CLAVATA1-related BAM1, BAM2 and BAM3 receptor kinase-like proteins are required for meristem function in *Arabidopsis*. Plant J *45*, 1–16.

[Donner et al., 2010] Donner, T. J., Sherr, I. and Scarpella, E. (2010). Auxin signal transduction in *Arabidopsis* vein formation. Plant Signal Behav *5*, 70–72.

[Earley et al., 2006] Earley, K. W., Haag, J. R., Pontes, O., Opper, K., Juehne, T., Song, K. and Pikaard, C. S. (2006). Gateway-compatible vectors for plant functional genomics and proteomics. Plant J *45*, 616–629.

[Edahiro and Seki, 2006] Edahiro, J. and Seki, M. (2006). Phenylpropanoid metabolite supports cell aggregate formation in strawberry cell suspension culture. J Biosci Bioeng *102*, 8–13.

[Edsall and Wyman, 1958] Edsall, T. and Wyman, J. (1958). Biophysical Chemistry. Academic Press, New York, p: 263-282.

[Ellis et al., 1996] Ellis, D. D., Zeldin, E. L., Brodhagen, M., Russin, W. A. and McCown, B. H. (1996). Taxol production in nodule cultures of *Taxus*. J Nat Prod *59*, 246–250.

[Endt et al., 2007] Endt, D. V., e Silva, M. S., Kijne, J. W., Pasquali, G. and Memelink, J. (2007). Identification of a bipartite jasmonate-responsive promoter element in the *Catharanthus roseus* ORCA3 transcription factor gene that interacts specifically with AT-Hook DNA-binding proteins. Plant Physiol *144*, 1680–1689.

[Endt et al., 2002] Endt, D. V., Kijne, J. W. and Memelink, J. (2002). Transcription factors controlling plant secondary metabolism: what regulates the regulators? Phytochemistry *61*, 107–114.

[Etchells and Turner, 2010] Etchells, J. P. and Turner, S. R. (2010). The PXY-CLE41 receptor ligand pair defines a multifunctional pathway that controls the rate and orientation of vascular cell division. Development *137*, 767–774.

[Eulgem et al., 2000] Eulgem, T., Rushton, P. J., Robatzek, S. and Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. Trends Plant Sci *5*, 199–206.

[Eulgem et al., 1999] Eulgem, T., Rushton, P. J., Schmelzer, E., Hahlbrock, K. and Somssich, I. E. (1999). Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors. EMBO J *18*, 4689–4699.

[Falkenstein et al., 1991] Falkenstein, E., Groth, B., Mithoefer, A. and Weiler, E. W. (1991). Methyljasmonate and alpha-linolenic acid are potent inducers of tendril coiling. PLANTA *185*, 316–322.

[Farmer and Ryan, 1990] Farmer, E. E. and Ryan, C. A. (1990). Interplant communication: airborne methyl jasmonate induces synthesis of proteinase inhibitors in plant leaves. Proc Natl Acad Sci U S A *87*, 7713–7716.

[Farnsworth, 1988] Farnsworth, N. (1988). Screening plants for new medicines. Chapter 9. National Academy Press, Washington D.C.

[Farrar and Donnison, 2007] Farrar, K. and Donnison, I. S. (2007). Construction and screening of BAC libraries made from *Brachypodium* genomic DNA. Nat Protoc *2*, 1661–1674.

[Fisher and Turner, 2007] Fisher, K. and Turner, S. (2007). PXY, a receptor-like kinase essential for maintaining polarity during plant vascular-tissue development. Curr Biol *17*, 1061–1066.

[Fuchs and Johnson, 1978] Fuchs, D. A. and Johnson, R. K. (1978). Cytologic evidence that taxol, an antineoplastic agent from *Taxus brevifolia*, acts as a mitotic spindle poison. Cancer Treat Rep *62*, 1219–1222.

[Fukuda, 1996] Fukuda, H. (1996). Xylogenesis: Initiation, progression, and cell death. Annu Rev Plant Physiol Plant Mol Biol *47*, 299–325.

[Fukuda, 1997] Fukuda, H. (1997). Tracheary element differentiation. Plant Cell *9*, 1147–1156.

[Fukuda, 2004] Fukuda, H. (2004). Signals that control plant vascular cell differentiation. Nat Rev Mol Cell Biol *5*, 379–391.

[Galis et al., 2006] Galis, I., Simek, P., Narisawa, T., Sasaki, M., Horiguchi, T., Fukuda, H. and Matsuoka, K. (2006). A novel R2R3 MYB transcription factor NtMYBJS1 is a methyl jasmonate-dependent regulator of phenylpropanoid-conjugate biosynthesis in tobacco. Plant J *46*, 573–592.

[Gallagher, 1992] Gallagher, S. R. (1992). GUS Protocols: Using the *GUS* Gene as a Reporter of Gene Expression. Academic Press, Boston.

[Gan, 1989] Gan, C. (1989). Gene Gun Accelerates DNA-Coated Particles To Transform Intact Cells. The Scientist *3*, 25.

[Gatehouse, 2002] Gatehouse, J. A. (2002). Plant resistance towards insect herbivores: a dynamic interaction. New Phytologist *156*, 145–169.

[Gershenzon et al., 1989] Gershenzon, J., Maffei, M. and Croteau, R. (1989). Biochemical and histochemical localization of monoterpene biosynthesis in the glandular trichomes of spearmint (*Mentha spicata*). Plant Physiol *89*, 1351–1357.

[Goodman and Walsh, 2001] Goodman, J. and Walsh, V. (2001). The story of taxol: nature and politics in the pursuit of an anti-cancer drug. Cambridge University Press.

[Goossens et al., 2003] Goossens, A., Haekkinen, S. T., Laakso, I., Seppaenen-Laakso, T., Biondi, S., Sutter, V. D., Lammertyn, F., Nuutila, A. M., Soederlund, H., Zabeau, M., Inze, D. and Oksman-Caldentey, K.-M. (2003). A functional genomics approach toward the understanding of secondary metabolism in plant cells. Proc Natl Acad Sci U S A *100*, 8595–8600.

[Gordon et al., 2007] Gordon, S. P., Heisler, M. G., Reddy, G. V., Ohno, C., Das, P. and Meyerowitz, E. M. (2007). Pattern formation during de novo assembly of the *Arabidopsis* shoot meristem. Development *134*, 3539–3548.

[Grafi and Avivi, 2004] Grafi, G. and Avivi, Y. (2004). Stem cells: a lesson from dedifferentiation. Trends Biotechnol *22*, 388–389.

[Green et al., 2006] Green, R. E., Krause, J., Ptak, S. E., Briggs, A. W., Ronan, M. T., Simons, J. F., D., L., Egholm, M., Rothberg, J. M., Paunovic, M. and Paeaebo, S. (2006). Analysis of one million base pairs of Neanderthal DNA. Nature *444*, 330–336.

[Grotewold et al., 1998] Grotewold, E., Chamberlin, M., Snook, M., Siame, B., Butler, L., Swenson, J., Maddock, S., Clair, G. S. and Bowen, B. (1998).

Engineering secondary metabolism in maize cells by ectopic expression of transcription factors. Plant Cell *10*, 721–740.

[Gu et al., 2002] Gu, Y.-Q., Wildermuth, M. C., Chakravarthy, S., Loh, Y.-T., Yang, C., He, X., Han, Y. and Martin, G. B. (2002). Tomato transcription factors pti4, pti5, and pti6 activate defense responses when expressed in *Arabidopsis*. Plant Cell *14*, 817–831.

[Gu et al., 2000] Gu, Y. Q., Yang, C., Thara, V. K., Zhou, J. and Martin, G. B. (2000). *Pti4* is induced by ethylene and salicylic acid, and its product is phosphorylated by the Pto kinase. Plant Cell *12*, 771–786.

[Guedes-Correa et al., 2008] Guedes-Correa, L. G., Riano-Pachon, D. M., Schrago, C. G., dos Santos, R. V., Mueller-Roeber, B. and Vincentz, M. (2008). The role of bZIP transcription factors in green plant evolution: adaptive features emerging from four founder genes. PLoS One *3*, e2944.

[Guerineau et al., 2003] Guerineau, F., Benjdia, M. and Zhou, D. X. (2003). A jasmonate-responsive element within the *A. thaliana VSP1* promoter. J Exp Bot *54*, 1153–1162.

[Guerineau et al., 1992] Guerineau, F., Lucy, A. and Mullineaux, P. (1992). Effect of two consensus sequences preceding the translation initiator codon on gene expression in plant protoplasts. Plant Mol Biol *18*, 815–818.

[Gundlach et al., 1992] Gundlach, H., Mueller, M. J., Kutchan, T. M. and Zenk, M. H. (1992). Jasmonic acid is a signal transducer in elicitor-induced plant cell cultures. Proc Natl Acad Sci U S A *89*, 2389–2393.

[Guo et al., 2006] Guo, B., Kai, G., Jin, H. and Tang, K. (2006). Taxol synthesis. African Journal of Biotechnology *5*, 15–20.

[Gurib-Fakim, 2006] Gurib-Fakim, A. (2006). Medicinal plants: traditions of yesterday and drugs of tomorrow. Mol Aspects Med *27*, 1–93.

[Hadlington and Denecke, 2001] Hadlington, J. L. and Denecke, J. (2001). Transient expression, a tool to adress questions in plant cell biology; In: Plant cell biology: a practical approach. University Press: Oxford, New York.

[Han et al., 1994] Han, K.-H., Fleming, P., Walker, K., Loper, M., Chilton, W. S., Mocek, U., Gordon, M. P. and Floss, H. G. (1994). Genetic transformation of mature *Taxus*: an approach to genetically control the in vitro production of the anticancer drug, taxol. Plant Science *95*, 187–196.

[Hara et al., 2000] Hara, K., Yagi, M., Kusano, T. and Sano, H. (2000). Rapid systemic accumulation of transcripts encoding a tobacco WRKY transcription factor upon wounding. Mol Gen Genet *263*, 30–37.

[Harborne, 1991] Harborne, J. B. (1991). Recent advances in the ecological chemistry of plant terpenoids. In Ecological chemistry and biochemistry of Plant Terpenoids. Claredon Press, Oxford.

[Hartman, 1996] Hartman, T. (1996). Diversity and variability of plant secondary metabolism: a mechanistic view. Entomologia Experimentalis et Applicata *80*, 177–188.

[Hartzell, 1991] Hartzell, H. (1991). The yew tree, a thousand whispers. Hulogosi Books.

[Harvey, 2007] Harvey, A. L. (2007). Natural products as a screening resource. Curr Opin Chem Biol *11*, 480–484.

[He and Gan, 2001] He, Y. and Gan, S. (2001). Identical promoter elements are involved in regulation of the *OPR1* gene by senescence and jasmonic acid in *Arabidopsis*. Plant Mol Biol *47*, 595–605.

[Hefner et al., 1998] Hefner, J., Ketchum, R. E. and Croteau, R. (1998). Cloning and functional expression of a cDNA encoding geranylgeranyl diphosphate synthase from *Taxus canadensis* and assessment of the role of this prenyltransferase in cells induced for taxol production. Arch Biochem Biophys *360*, 62–74.

[Heim et al., 2003] Heim, M. A., Jakoby, M., Werber, M., Martin, C., Weisshaar, B. and Bailey, P. C. (2003). The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. Mol Biol Evol *20*, 735–747.

[Hein et al., 2005] Hein, I., Williamson, S., Russell, J. and Powell, W. (2005). Isolation of high molecular weight DNA suitable for BAC library construction from woody perennial soft-fruit species. Biotechniques *38*, 69–71.

[Hellens et al., 2000] Hellens, R. P., Edwards, E. A., Leyland, N. R., Bean, S. and Mullineaux, P. M. (2000). pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. Plant Mol Biol *42*, 819–832.

[Herms and Mattson, 1992] Herms, D. and Mattson, W. J. (1992). The dilemma of plants: to grow or defend. Q Rev Biol *67*, 283–335.

[Hilson et al., 2004] Hilson, P., Allemeersch, J., Altmann, T., Aubourg, S., Avon, A., Beynon, J., Bhalerao, R. P., Bitton, F., Caboche, M., Cannoot, B., Chardakov, V., Cognet-Holliger, C., Colot, V., Crowe, M., Darimont, C., Durinck, S., Eickhoff, H., de Longevialle, A. F., Farmer, E. E., Grant, M., Kuiper, M. T. R., Lehrach, H., Leon, C., Leyva, A., Lundeberg, J., Lurin, C., Moreau, Y., Nietfeld, W., Paz-Ares, J., Reymond, P., Rouze, P., Sandberg, G., Segura, M. D., Serizet, C., Tabrett, A., Taconnat, L., Thareau, V., van Hummelen, P., Vercruysse, S., Vuylsteke, M., Weingartner, M., Weisbeek, P. J., Wirta, V., Wittink, F. R. A., Zabeau, M. and Small, I. (2004). Versatile gene-specific sequence tags for *Arabidopsis* functional genomics: transcript profiling and reverse genetics applications. Genome Res *14*, 2176–2189.

[Hoff, 1998] Hoff, D. D. V. (1998). Phase I trials of dexrazoxane and other potential applications for the agent. Semin Oncol *25*, 31–36.

[Holmes et al., 1991] Holmes, F. A., Walters, R. S., Theriault, R. L., Forman, A. D., Newton, L. K., Raber, M. N., Buzdar, A. U., Frye, D. K. and Hortobagyi, G. N. (1991). Phase II trial of taxol, an active drug in the treatment of metastatic breast cancer. J Natl Cancer Inst *83*, 1797–1805.

[Horowitz et al., 1979] Horowitz, S., Schiff, P. and Fant, J. (1979). Promotion of microtubule assembly in vitro by taxol. Nature *277*, 665–667.

[Huang et al., 1986] Huang, C. H., Kingston, D. G., Magri, N. F., Samaranayake, G. and Boettner, F. E. (1986). New taxanes from *Taxus brevifolia*, 2. J Nat Prod *49*, 665–669.

[Hulst et al., 1989] Hulst, A. C., Meyer, M. M. T., Breteler, H. and Tramper, J. (1989). Effect of aggregate size in cell cultures of *Tagetes patula* on thiophene production and cell growth. Applied Microbiology and Biotechnology *30*, 18–25.

[Hurst, 1995] Hurst, H. C. (1995). Transcription factors 1: bZIP proteins. Protein Profile *2*, 101–168.

[Hyams and Lloyd, 1994] Hyams, J. S. and Lloyd, C. W. (1994). Microtubules. Wiley-Liss, cop. New York ; Chichester ; Brisbane.

[Ibanes et al., 2009] Ibanes, M., Fabregas, N., Chory, J. and Cano-Delgado, A. I. (2009). Brassinosteroid signaling and auxin transport are required to establish the periodic pattern of *Arabidopsis* shoot vascular bundles. Proc Natl Acad Sci U S A *106*, 13630–13635.

[Initiative, 2000] Initiative, A. G. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature *408*, 796–815.

[Ishiguro et al., 2001] Ishiguro, S., Kawai-Oda, A., Ueda, J., Nishida, I. and Okada, K. (2001). The *DEFECTIVE IN ANTHER DEHISCIENCE* gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in *Arabidopsis*. Plant Cell *13*, 2191–2209.

[Jacrot et al., 1983] Jacrot, M., Riondel, J., Picot, F., Leroux, D., Mouriquand, C., Beriel, H. and Potier, P. (1983). Action of taxol on human tumors transplanted in athymic mice. C R Seances Acad Sci III *297*, 597–600.

[Jander et al., 2001] Jander, G., Cui, J., Nhan, B., Pierce, N. E. and Ausubel, F. M. (2001). The *TASTY* locus on chromosome 1 of *Arabidopsis* affects feeding of the insect herbivore *Trichoplusia ni*. Plant Physiol *126*, 890–898.

[Jaziri et al., 1996] Jaziri, M., Zhiri, A., Guo, Y.-W., Dupont, J.-P., Shimomura, K., Hamada, H., Vanhaelen, M. and Homes, J. (1996). *Taxus sp.* cell, tissue and organ cultures as alternative sources for taxoids production: a literature survey. Plant Cell, Tissue and Organ Culture *46*, 59–75.

[Jefferson, 1989] Jefferson, R. A. (1989). The *GUS* reporter gene system. Nature *342*, 837–838.

[Jemal et al., 2007] Jemal, A., Siegel, R., Ward, E., Murray, T., X., J. and Thun, M. J. (2007). Cancer statistics, 2007. CA Cancer J Clin *57*, 43–66.

[Jennewein et al., 2004] Jennewein, S., Long, R. M., Williams, R. M. and Croteau, R. (2004). Cytochrome p450 taxadiene 5alpha-hydroxylase, a mechanistically unusual monooxygenase catalyzing the first oxygenation step of taxol biosynthesis. Chem Biol *11*, 379–387.

[Jennewein et al., 2003] Jennewein, S., Rithner, C. D., Williams, R. M. and Croteau, R. (2003). Taxoid metabolism: Taxoid 14beta-hydroxylase is a cytochrome P450-dependent monooxygenase. Arch Biochem Biophys *413*, 262–270.

[Jennewein et al., 2001] Jennewein, S., Rithner, C. D., Williams, R. M. and Croteau, R. B. (2001). Taxol biosynthesis: taxane 13 alpha-hydroxylase is a cytochrome P450-dependent monooxygenase. Proc Natl Acad Sci U S A *98*, 13595–13600.

[Jennewein et al., 2004] Jennewein, S., Wildung, M. R., Chau, M., Walker, K. and Croteau, R. (2004). Random sequencing of an induced *Taxus* cell cDNA library for identification of clones involved in Taxol biosynthesis. Proc Natl Acad Sci U S A *101*, 9149–9154.

[Jin and Martin, 1999] Jin, H. and Martin, C. (1999). Multifunctionality and diversity within the plant *MYB*-gene family. Plant Mol Biol *41*, 577–585.

[Johri and Bhojwani, 1965] Johri, M. B. and Bhojwani, S. S. (1965). Growth responses of mature endosperm in cultures. Nature *208*, 1345–1347.

[Kang et al., 2010] Kang, B., Wang, H., Nam, K. H., Li, J. and Li, J. (2010). Activation-tagged suppressors of a weak brassinosteroid receptor mutant. Mol Plant *3*, 260–268.

[Kato et al., 2007] Kato, N., Dubouzet, E., Kokabu, Y., Yoshida, S., Taniguchi, Y., Dubouzet, J. G., Yazaki, K. and Sato, F. (2007). Identification of a WRKY protein as a transcriptional regulator of benzylisoquinoline alkaloid biosynthesis in *Coptis japonica*. Plant Cell Physiol *48*, 8–18.

[Katterman and Shattuck, 1983] Katterman, F. R. and Shattuck, V. I. (1983). An effective method of DNA isolation from the mature leaves of *Gossypium* species that contain large amounts of phenolic terpenoids and tannins. Prep Biochem *13*, 347–359.

[Keller-Juslen et al., 1971] Keller-Juslen, C., Kuhn, M. and von Wartburg, A. (1971). Synthesis and antimitotic activity of glycosidic lignan derivates related to podophyllotoxin. Journal of Medicinal Chemistry *14*, 936–940.

[Ketchum et al., 1999] Ketchum, R. E., Gibson, D. M., Croteau, R. B. and Shuler, M. L. (1999). The kinetics of taxoid accumulation in cell suspension cultures of *Taxus* following elicitation with methyl jasmonate. Biotechnol Bioeng *62*, 97–105.

[Ketchum et al., 2007] Ketchum, R. E. B., Wherland, L. and Croteau, R. B. (2007). Stable transformation and long-term maintenance of transgenic *Taxus* cell suspension cultures. Plant Cell Rep *26*, 1025–1033.

[Kiger et al., 2001] Kiger, A. A., Jones, D. L., Schulz, C., Rogers, M. B. and Fuller, M. T. (2001). Stem cell self-renewal specified by JAK-STAT activation in response to a support cell cue. Science *294*, 2542–2545.

[Kim et al., 1997] Kim, C. S., Lee, C. H., Shin, J. S., Chung, Y. S. and Hyung, N. I. (1997). A simple and rapid method for isolation of high quality genomic DNA from fruit trees and conifers using PVP. Nucleic Acids Res *25*, 1085–1086.

[Kim et al., 1993] Kim, S. R., Kim, Y. and An, G. (1993). Identification of methyl jasmonate and salicylic acid response elements from the *NOPALINE SYNTHASE* (*NOS*) promoter. Plant Physiol *103*, 97–103.

[Kim et al., 1992] Kim, S. R., Kim, Y., Costa, M. A. and An, G. (1992). Inhibition of sucrose enhancer effect of the potato proteinase Inhibitor II promoter by salicylic acid. Plant Physiol *98*, 1479–1483.

[Kingsbury et al., 1991] Kingsbury, W. D., Boehm, J. C., Jakas, D. R., Holden, K. G., Hecht, S. M., Gallagher, G., Caranfa, M. J., McCabe, F. L., Faucette, L. F. and Johnson, R. K. (1991). Synthesis of water-soluble (aminoalkyl) camptothecin analogues: inhibition of topoisomerase I and antitumor activity. J Med Chem *34*, 98–107.

[Kingston et al., 1982] Kingston, D. G., Hawkins, D. R. and Ovington, L. (1982). New taxanes from *Taxus brevifolia*. J Nat Prod *45*, 466–470.

[Klein et al., 1992] Klein, T. M., Arentzen, R., Lewis, P. A. and Fitzpatrick-McElligott, S. (1992). Transformation of microbes, plants and animals by particle bombardment. Biotechnology (N Y) *10*, 286–291.

[Kleining, 1989] Kleining, H. (1989). The role of plastids in isoprenoid biosynthesis. Annual Review of Plant Physiology *40*, 39–59.

[Klemm et al., 1998] Klemm, J. D., Schreiber, S. L. and Crabtree, G. R. (1998). Dimerization as a regulatory mechanism in signal transduction. Annu Rev Immunol *16*, 569–592.

[Krishnaswamy et al., 2011] Krishnaswamy, S., Verma, S., Rahman, M. H. and Kav, N. N. V. (2011). Functional characterization of four *APETALA2*-family genes (*RAP2.6, RAP2.6L, DREB19 and DREB26*) in *Arabidopsis*. Plant Mol Biol *75*, 107–127.

[Kunimoto et al., 1987] Kunimoto, T., Nitta, K., Tanaka, T., Uehara, N., Baba, H., Takeuchi, M., Yokokura, T., Sawada, S., Miyasaka, T. and Mutai, M. (1987). Antitumor activity of 7-ethyl-10-[4-(1-piperidino)-1-

piperidino]carbonyloxy-camptothec in, a novel water-soluble derivative of camptothecin, against murine tumors. Cancer Res *47*, 5944–5947.

[Latchman, 1997] Latchman, D. S. (1997). Transcription factors: an overview. Int J Biochem Cell Biol *29*, 1305–1312.

[Lee et al., 2010] Lee, E.-K., Jin, Y.-W., Park, J. H., Yoo, Y. M., Hong, S. M., Amir, R., Yan, Z., Kwon, E., Elfick, A., Tomlinson, S., Halbritter, F., WAIBEL, T., Yun, B.-W. and Loake, G. J. (2010). Cultured cambial meristematic cells as a source of plant natural products. Nat Biotechnol *28*, 1213–1217.

[Leistner, 2005] Leistner, E. (2005). Die Biologie der Taxane, Arzneimittel aus der Natur. Pharm. Unserer Zeit *34 (2)*, 98.

[Levy et al., 2005] Levy, M., Rachmilevitch, S. and Abel, S. (2005). Transient *Agrobacterium*-mediated gene expression in the *Arabidopsis* hydroponics root system for subcellular localization studies. Plant Molecular Biology Reporter *23*, 179–184.

[Li et al., 2004] Li, L., Zhao, Y., McCaig, B. C., Wingerd, B. A., Wang, J., Whalon, M. E., Pichersky, E. and Howe, G. A. (2004). The tomato homolog of CORONATINE-INSENSITIVE1 is required for the maternal control of seed maturation, jasmonate-signaled defense responses, and glandular trichome development. Plant Cell *16*, 126–143.

[Libault et al., 2007] Libault, M., Wan, J., Czechowski, T., Udvardi, M. and Stacey, G. (2007). Identification of 118 *Arabidopsis* transcription factor and

30 ubiquitin-ligase genes responding to chitin, a plant-defense elicitor. Mol Plant Microbe Interact *20*, 900–911.

[Lindholm, 2005] Lindholm, P. (2005). Cytotoxic compounds of plant origin - Biological and chemical diversity. PhD thesis, University of Uppsala.

[Lopato et al., 2006] Lopato, S., Bazanova, N., Morran, S., Milligan, A. S., Shirley, N. and Langridge, P. (2006). Isolation of plant transcription factors using a modified yeast one-hybrid system. Plant Methods *2*, 3.

[Lorenzo et al., 2004] Lorenzo, O., Chico, J. M., Sanchez-Serrano, J. J. and Solano, R. (2004). *JASMONATE-INSENSITIVE1* encodes a MYC transcription factor essential to discriminate between different jasmonate-regulated defense responses in *Arabidopsis*. Plant Cell *16*, 1938–1950.

[Lorenzo et al., 2003] Lorenzo, O., Piqueras, R., Sanchez-Serrano, J. J. and Solano, R. (2003). ETHYLENE RESPONSE FACTOR1 integrates signals from ethylene and jasmonate pathways in plant defense. Plant Cell *15*, 165–178.

[Lorenzo and Solano, 2005] Lorenzo, O. and Solano, R. (2005). Molecular players regulating the jasmonate signalling network. Curr Opin Plant Biol *8*, 532–540.

[Madhusudhan and Ravishankar, 1996] Madhusudhan, R. and Ravishankar, G. A. (1996). Gradient of anthocyanin in cell aggregates of *Daucus carota* in suspension cultures. Biotechnology Letters *18*, 1253–1256.

[Maehoenen et al., 2000] Maehoenen, A. P., Bonke, M., Kauppinen, L., Riikonen, M., Benfey, P. N. and Helariutta, Y. (2000). A novel two-component hybrid molecule regulates vascular morphogenesis of the *Arabidopsis* root. Genes Dev *14*, 2938–2943.

[Mahalingam et al., 2003] Mahalingam, R., Gomez-Buitrago, A., Eckardt, N., Shah, N., Guevara-Garcia, A., Day, P., Raina, R. and Fedoroff, N. V. (2003). Characterizing the stress/defense transcriptome of *Arabidopsis*. Genome Biol *4*, R20.

[Mahmoud and Croteau, 2001] Mahmoud, S. S. and Croteau, R. B. (2001). Metabolic engineering of essential oil yield and composition in mint by altering expression of deoxyxylulose phosphate reductoisomerase and menthofuran synthase. Proc Natl Acad Sci U S A *98*, 8915–8920.

[Mangatal et al., 1989] Mangatal, L., Adeline, M.-T., Guenard, D., Gueritte-Voegelein, D. and Potier, P. (1989). Application of the vicinal oxymination reaction with asymmetric induction to the hemisynthesis of taxol and analogues. Tetrahedron.

[Mangeney et al., 1979] Mangeney, P., Andriamialisoa, R., Lallemand, J., Langlois, N., Langlois, Y. and Potier, P. (1979). 5-nor anhydrovinblastine: Prototype of a new class of vinblastine derivates. Tetrahydron, Stuttgart, Germany.

[Mardis, 2008] Mardis, E. R. (2008). Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet *9*, 387–402.

[Margulies et al., 2005] Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L. I., Jarvie, T. P., Jirage, K. B., Kim, J.-B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F. and Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature *437*, 376–380.

[Mason et al., 1993] Mason, H. S., DeWald, D. B. and Mullet, J. E. (1993). Identification of a methyl jasmonate-responsive domain in the soybean vspB promoter. Plant Cell *5*, 241–251.

[Maury et al., 1999] Maury, S., Geoffroy, P. and Legrand, M. (1999). Tobacco O-methyltransferases involved in phenylpropanoid metabolism. The different caffeoyl-coenzyme A/5-hydroxyferuloyl-coenzyme A 3/5-O-methyltransferase and caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase classes have distinct substrate specificities and expression patterns. Plant Physiol *121*, 215–224.

[McConn et al., 1997] McConn, M., Creelman, R. A., Bell, E., Mullet, J. E. and Browse, J. (1997). Jasmonate is essential for insect defense in *Arabidop-*

sis. Proc Natl Acad Sci U S A *94*, 5473–5477.

[McGarvey and Croteau, 1995] McGarvey, D. J. and Croteau, R. (1995). Terpenoid metabolism. Plant Cell *7*, 1015–1026.

[McLaughlin et al., 1981] McLaughlin, J. L., Miller, R. W., Powell, R. G. and Smith, C. R. (1981). 19-Hydroxybaccatin III, 10-deacetylcephalomannine, and 10-deacetyltaxol: new antitumor taxanes from *Taxus wallichiana*. J Nat Prod *44*, 312–319.

[Memelink et al., 2001] Memelink, J., Verpoorte, R. and Kijne, J. W. (2001). ORCAnization of jasmonate-responsive gene expression in alkaloid metabolism. Trends Plant Sci *6*, 212–219.

[Menke et al., 1999] Menke, F. L., Champion, A., Kijne, J. W. and Memelink, J. (1999). A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene Str interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2. EMBO J *18*, 4455–4463.

[Michiels et al., 2003] Michiels, A., van den Ende, W., Tucker, M., van Riet, L. and van Laere, A. (2003). Extraction of high-quality genomic DNA from latex-containing plants. Anal Biochem *315*, 85–89.

[Mirjalili and Linden, 1996] Mirjalili, N. and Linden, J. C. (1996). Methyl jasmonate induced production of taxol in suspension cultures of *Taxus cuspidata*: ethylene interaction and induction models. Biotechnol Prog *12*, 110–118.

[Mol et al., 1998] Mol, J., Grotewoldb, E. and Koes, R. (1998). How genes paint flowers and seeds. Trends in Pland Science *3*, 212–217.

[Morozova and Marra, 2008] Morozova, O. and Marra, M. A. (2008). Applications of next-generation sequencing technologies in functional genomics. Genomics *92*, 255–264.

[Morrissey, 2003] Morrissey, S. R. (2003). Maximising Returns, A driving force in getting Taxol to market quickly, NIH is being asked to evaluate the way it deals with the private-sector partnerships. CENEAR *81*, 17–20.

[Mukherjee et al., 2002] Mukherjee, S., Ghosh, B., Jha, T. B. and Jha, S. (2002). Variation in content of taxol and related taxanes in Eastern Himalayan populations of *Taxus wallichiana*. Planta Med *68*, 757–759.

[Mumm and Hilker, 2006] Mumm, R. and Hilker, M. (2006). Direct and indirect chemical defence of pine against folivorous insects. Trends Plant Sci *11*, 351–358.

[Murray, 1998] Murray, B. G. (1998). Nuclear DNA amounts in gymnosperms. ANNALS OF BOTANY *82*, 3–15.

[Nagata et al., 2001] Nagata, N., Asami, T. and Yoshida, S. (2001). Brassinazole, an inhibitor of brassinosteroid biosynthesis, inhibits development of secondary xylem in cress plants (*Lepidium sativum*). Plant Cell Physiol *42*, 1006–1011.

[Naill and Roberts, 2005a] Naill, M. C. and Roberts, S. C. (2005a). Cell cycle analysis of *Taxus* suspension cultures at the single cell level as an indicator of culture heterogeneity. Biotechnol Bioeng *90*, 491–500.

[Naill and Roberts, 2005b] Naill, M. C. and Roberts, S. C. (2005b). Flow cytometric analysis of protein content in *Taxus* protoplasts and single cells as compared to aggregated suspension cultures. Plant Cell Rep *23*, 528–533.

[Nakatsukaa et al., 2009] Nakatsukaa, A., Yamagishib, M., Nakanoa, M., Tasakia, K. and Kobayashia, N. (2009). Light-induced expression of basic helix-loop-helix genes involved in anthocyanin biosynthesis in flowers and leaves of *Asiatic* hybrid lily. Scientia Horticulturae *121*, 84–91.

[Nam et al., 2003] Nam, J., dePamphilis, C. W., Ma, H. and Nei, M. (2003). Antiquity and evolution of the *MADS-box* gene family controlling flower development in plants. Mol Biol Evol *20*, 1435–1447.

[Narvaez-Vasqueza et al., 1999] Narvaez-Vasqueza, J., Florin-Christensenb, J. and Ryana, C. A. (1999). Positional specificity of a phospholipase A activity induced by wounding, systemin, and oligosaccharide elicitors in tomato leaves. Plant Cell *11*, 2249–2260.

[Nater et al., 2005] Nater, U. M., Rohleder, N., Gaab, J., Berger, S., Jud, A., Kirschbaum, C. and Ehlert, U. (2005). Human salivary alpha-amylase reactivity in a psychosocial stress paradigm. Int J Psychophysiol *55*, 333–342.

[Neuss et al., 1962] Neuss, N., Gorman, M., Boaz, H. and Cone, N. (1962). Vinca alkaloids, XI. Structures of leurocristine (LCR) and vincaleukoblastine (VLB). Journal of the American Chemical Society *84*, 1509–1510.

[Neuss et al., 1964] Neuss, N., Gorman, M., Hargrove, W., Cone, N., Biemann, K., Buechi, G. and Manning, R. (1964). Vinca alkaloids XXI. The structures of the oncolytic alkaloids vinblastine (VBL) and vincristine (VCR). Journal of the American Chemical Society *86*, 1440–1442.

[Neuss et al., 1959] Neuss, N., Gorman, M., Svoboda, G., Maciak, G. and Beer, C. (1959). Vinca alkaloids III. Characterization of leurosine and vincaleukoblastine, new alkaloids from *Vinca rosea* Linn. Journal of the American Chemical Society *81*, 4754–4755.

[Newman et al., 2003] Newman, D. J., Cragg, G. M. and Snader, K. M. (2003). Natural products as sources of new drugs over the period 1981-2002. J Nat Prod *66*, 1022–1037.

[Nims et al., 2006] Nims, E., Dubois, C. P., Roberts, S. C. and Walker, E. L. (2006). Expression profiling of genes involved in paclitaxel biosynthesis for targeted metabolic engineering. Metab Eng *8*, 385–394.

[Noonan et al., 2006] Noonan, J. P., Coop, G., Kudaravalli, S., Smith, D., Krause, J., Alessi, J., Chen, F., Platt, D., Paeaebo, S., Pritchard, J. K. and Rubin, E. M. (2006). Sequencing and analysis of Neanderthal genomic DNA. Science *314*, 1113–1118.

[Oberlies and Kroll, 2004] Oberlies, N. H. and Kroll, D. J. (2004). Camptothecin and taxol: historic achievements in natural products research. J Nat Prod *67*, 129–135.

[Ohme-Takagi and Shinshi, 1995] Ohme-Takagi, M. and Shinshi, H. (1995). Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. Plant Cell *7*, 173–182.

[Ohwi, 1965] Ohwi, J. (1965). Flora of Japan. Smithsonian Institute Press. Wshington D.C.

[Okouneva et al., 2003] Okouneva, T., Hill, B. T., Wilson, L. and Jordan, M. A. (2003). The effects of vinflunine, vinorelbine, and vinblastine on centromere dynamics. Mol Cancer Ther *2*, 427–436.

[Osbourn, 2010] Osbourn, A. (2010). Gene clusters for secondary metabolic pathways: an emerging theme in plant biology. Plant Physiol *154*, 531–535.

[Parchman et al., 2010] Parchman, T. L., Geist, K. S., Grahnen, J. A., Benkman, C. W. and Buerkle, C. A. (2010). Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. BMC Genomics *11*, 180.

[Parkin et al., 2001] Parkin, D. M., Bray, F., Ferlay, J. and Pisani, P. (2001). Estimating the world cancer burden: Globocan 2000. Int J Cancer *94*, 153–156.

[Paschold et al., 2007] Paschold, A., Halitschke, R. and Baldwin, I. T. (2007). Co(i)-ordinating defenses: NaCOI1 mediates herbivore- induced resistance

in *Nicotiana attenuata* and reveals the role of herbivore movement in avoiding defenses. Plant J *51*, 79–91.

[Pauwels et al., 2010] Pauwels, L., Barbero, G. F., Geerinck, J., Tilleman, S., Grunewald, W., Perez, A. C., Chico, J. M., Bossche, R. V., Sewell, J., Gil, E., Garcia-Casado, G., Witters, E., Inze, D., Long, J. A., Jaeger, G. D., Solano, R. and Goossens, A. (2010). NINJA connects the co-repressor TOPLESS to jasmonate signalling. Nature *464*, 788–791.

[Pauwels et al., 2009] Pauwels, L., Inze, D. and Goossens, A. (2009). Jasmonate-inducible gene: What does it mean? Trends Plant Sci *14*, 87–91.

[Pauwels et al., 2008] Pauwels, L., Morreel, K., Witte, E. D., Lammertyn, F., Montagu, M. V., Boerjan, W., Inze, D. and Goossens, A. (2008). Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured *Arabidopsis* cells. Proc Natl Acad Sci U S A *105*, 1380–1385.

[Pelletier et al., 1999] Pelletier, M. K., Burbulis, I. E. and Winkel-Shirley, B. (1999). Disruption of specific flavonoid genes enhances the accumulation of flavonoid enzymes and end-products in *Arabidopsis* seedlings. Plant Mol Biol *40*, 45–54.

[Pelta et al., 1996] Pelta, J., Livolant, F. and Sikorav, J. L. (1996). DNA aggregation induced by polyamines and cobalthexamine. J Biol Chem *271*, 5656–5662.

[Penninckx et al., 1998] Penninckx, I. A., Thomma, B. P., Buchala, A., Metraux, J. P. and Broekaert, W. F. (1998). Concomitant activation of jasmonate and ethylene response pathways is required for induction of a plant defensin gene in *Arabidopsis*. Plant Cell *10*, 2103–2113.

[Pepin et al., 1999] Pepin, M. F., Smith, M. and Reid, J. (1999). Application of imaging tools to plant cell culture: relationship between plant cell aggregation and flavonoid production. In Vitro Cellular and Developmental Biology - Plant *30*, 290–295.

[Peterson et al., 2002] Peterson, D. G., Tomkins, J. P., Frisch, D. A., Wing, R. A. and Paterson, A. H. (2002). Construction of plant bacterial artificial chromosome (BAC) libraries: An illustrated Guide.

[Phillips et al., 1999] Phillips, M. A., Savage, T. J. and Croteau, R. (1999). Monoterpene synthases of loblolly pine (*Pinus taeda*) produce pinene isomers and enantiomers. Arch Biochem Biophys *372*, 197–204.

[Pichersky and Gang, 2000] Pichersky, E. and Gang, D. R. (2000). Genetic and biochemistry of secondary metabolites in plants: an evolutionary perspective. Trends in Pland Science *5*, 439–445.

[Pichersky et al., 1994] Pichersky, E., Raguso, R. A., Lewinsohn, E. and Croteau, R. (1994). Floral scent production in *Clarkia* (Onagraceae) (I. localization and developmental modulation of monoterpene emission and Linalool Synthase activity). Plant Physiol *106*, 1533–1540.

[Pluthero, 1993] Pluthero, F. G. (1993). Rapid purification of high-activity Taq DNA polymerase. Nucleic Acids Res *21*, 4850–4851.

[Pommier, 1993] Pommier, Y. (1993). DNA topoisomerase I and II in cancer chemotherapy: update and perspectives. Cancer Chemother Pharmacol *32*, 103–108.

[Porebski et al., 1997] Porebski, S., Bailey, G. and Baum, R. (1997). Modification of CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol. Biol. Rep. *15*, 8–15.

[Pre et al., 2000] Pre, M., Siberil, Y., Memelink, J., Champion, A., Doireau, P. and Gantet, P. (2000). Isolation by the yeast one-hybrid system of cDNAs encoding transcription factors that bind to the G-box element of the strictosidine synthase gene promoter from *Catharanthus roseus*. Int J Bio-chrom *5*, 229–244.

[Qiu et al., 2009] Qiu, D., Pan, X., Wilson, I. W., Li, F., Liu, M., Teng, W. and Zhang, B. (2009). High throughput sequencing technology reveals that the taxoid elicitor methyl jasmonate regulates microRNA expression in Chinese yew (*Taxus chinensis*). Gene *436*, 37–44.

[Quattrocchio et al., 1998] Quattrocchio, F., Wing, J. F., van der Woude, K., Mol, J. N. and Koes, R. (1998). Analysis of bHLH and MYB domain proteins: species-specific regulatory differences are caused by divergent evolution of target anthocyanin genes. Plant J *13*, 475–488.

[Ramalho-Santos et al., 2002] Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R. C. and Melton, D. A. (2002). "Stemness": transcriptional profiling of embryonic and adult stem cells. Science *298*, 597–600.

[Rates, 2001] Rates, S. M. (2001). Plants as source of drugs. Toxicon *39*, 603–613.

[Reya et al., 2001] Reya, T., Morrison, S. J., Clarke, M. F. and Weissman, I. L. (2001). Stem cells, cancer, and cancer stem cells. Nature *414*, 105–111.

[Reymond and Farmer, 1998] Reymond, P. and Farmer, E. E. (1998). Jasmonate and salicylate as global signals for defense gene expression. Curr Opin Plant Biol *1*, 404–411.

[Rhoades, 1979] Rhoades, D. (1979). Evolution of plant chemical defense against herbivores. In: Herbivores: Their Interaction with Secondary Plant Metabolites. Academic Press, New York.

[Riechmann and Meyerowitz, 1998] Riechmann, J. L. and Meyerowitz, E. M. (1998). The AP2/EREBP family of plant transcription factors. Biol Chem *379*, 633–646.

[Riondel et al., 1986] Riondel, J., Jacrot, M., Picot, F., Beriel, H., Mouriquand, C. and Potier, P. (1986). Therapeutic response to taxol of six human tumors xenografted into nude mice. Cancer Chemother Pharmacol *17*, 137–142.

[Roberts, 2007] Roberts, S. C. (2007). Production and engineering of terpenoids in plant cell culture. Nat Chem Biol *3*, 387–395.

[Robinson et al., 2010] Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

[Romero et al., 1998] Romero, I., Fuertes, A., Benito, M. J., Malpica, J. M., Leyva, A. and Paz-Ares, J. (1998). More than 80R2R3-MYB regulatory genes in the genome of *Arabidopsis thaliana*. Plant J *14*, 273–284.

[Rothberg and Leamon, 2008] Rothberg, J. M. and Leamon, J. H. (2008). The development and impact of 454 sequencing. Nat Biotechnol *26*, 1117–1124.

[Rouster et al., 1997] Rouster, J., Leah, R., Mundy, J. and Cameron-Mills, V. (1997). Identification of a methyl jasmonate-responsive region in the promoter of a lipoxygenase 1 gene expressed in barley grain. Plant J *11*, 513–523.

[Rushton et al., 1995] Rushton, P. J., Macdonald, H., Huttly, A. K., Lazarus, C. M. and Hooley, R. (1995). Members of a new family of DNA-binding proteins bind to a conserved cis-element in the promoters of alpha-Amy2 genes. Plant Mol Biol *29*, 691–702.

[Rushton and Somssich, 1998] Rushton, P. J. and Somssich, I. E. (1998). Transcriptional control of plant genes responsive to pathogens. Curr Opin Plant Biol *1*, 311–315.

[Rushton et al., 1996] Rushton, P. J., Torres, J. T., Parniske, M., Wernert, P., Hahlbrock, K. and Somssich, I. E. (1996). Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley *PR1* genes. EMBO J *15*, 5690–5700.

[Russin et al., 1995] Russin, W., Ellis, D., Gottwald, J., Zelden, E., Brodhagen, M. and Evert, F. (1995). Immunocytochemical localization of taxol in *Taxus cuspidata*. Int. J. Plant Sci *156*, 668–678.

[Sambrook et al., 1989] Sambrook, J., Fritish, E. and Manaiatis, T. (1989). Molecular Cloning - A Laboratory Manual. Cold Spring Harbor Laboratory Press, USA.

[Sasaki-Sekimoto et al., 2005] Sasaki-Sekimoto, Y., Taki, N., Obayashi, T., Aono, M., Matsumoto, F., Sakurai, N., Suzuki, H., Hirai, M. Y., Noji, M., Saito, K., Masuda, T., ichiro Takamiya, K., Shibata, D. and Ohta, H. (2005). Coordinated activation of metabolic pathways for antioxidants and defence compounds by jasmonates and their roles in stress tolerance in *Arabidopsis*. Plant J *44*, 653–668.

[Saunders et al., 1995] Saunders, J. A., Lin, C. H., Hou, B. H., Cheng, J., Tsengwa, N., Lin, J. J., Smith, C. R., McIntosh, M. S. and Wert, S. V. (1995). Rapid optimization of electroporation conditions for plant cells, protoplasts, and pollen. Mol Biotechnol *3*, 181–190.

[Scheres et al., 1995] Scheres, B., Laurenzio, L. D., Willemsen, V., Hauser, M.-T., Janmaat, K., Weisbeek, P. and Benfey, P. N. (1995). Mutations affecting the radial organisation of the *Arabidopsis* root display specific defects throughout the embryonic axis. Development *121*, 53–62.

[Schmid and Blaxter, 2008] Schmid, R. and Blaxter, M. L. (2008). annot8r: GO, EC and KEGG annotation of EST datasets. BMC Bioinformatics *9*, 180.

[Schoendorf et al., 2001] Schoendorf, A., Rithner, C. D., Williams, R. M. and Croteau, R. B. (2001). Molecular cloning of a cytochrome P450 taxane 10

beta-hydroxylase cDNA from *Taxus* and functional expression in yeast. Proc Natl Acad Sci U S A *98*, 1501–1506.

[Schuetze et al., 2008] Schuetze, K., Harter, K. and Chaban, C. (2008). Post-translational regulation of plant bZIP factors. Trends Plant Sci *13*, 247–255.

[Schwartz and Cantor, 1984] Schwartz, D. C. and Cantor, C. R. (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell *37*, 67–75.

[Sehr et al., 2010] Sehr, E. M., Agusti, J., Lehner, R., Farmer, E. E., Schwarz, M. and Greb, T. (2010). Analysis of secondary growth in the *Arabidopsis* shoot reveals a positive role of jasmonate signalling in cambium formation. Plant J *63*, 811–822.

[Selth et al., 2010] Selth, L. A., Sigurdsson, S. and Svejstrup, J. Q. (2010). Transcript elongation by RNA polymerase II. Annu Rev Biochem *79*, 271–293.

[Sembdner and Parthier, 1993] Sembdner, G. and Parthier, B. (1993). The biochemistry and the physiological and molecular actions of jasmonates. Plant Molecular Biology *44*, 569–589.

[Sheen, 2001] Sheen, J. (2001). Signal transduction in maize and *Arabidopsis* mesophyll protoplasts. Plant Physiol *127*, 1466–1475.

[Shigyo et al., 2006] Shigyo, M., Hasebe, M. and Ito, M. (2006). Molecular evolution of the AP2 subfamily. Gene *366*, 256–265.

[Siebert et al., 1995] Siebert, P. D., Chenchik, A., Kellogg, D. E., Lukyanov, K. A. and Lukanov, S. A. (1995). An improved PCR method for walking in uncloned genomic DNA. Nucleic Acids Res *23*, 1087–1088.

[Singh, 1998] Singh, K. B. (1998). Transcriptional regulation in plants: the importance of combinatorial control. Plant Physiol *118*, 1111–1120.

[Singh and Bhalla, 2006] Singh, M. B. and Bhalla, P. L. (2006). Plant stem cells carve their own niche. Trends Plant Sci *11*, 241–246.

[Sitte et al., 1991] Sitte, P., Ziegler, H., Ehrendorfer, F. and Bresinsky, A. (1991). Strasburger Lehrbuch der Botanik. Gustav Fisher Verlag: Stuttgart, Jena, New York.

[Smet et al., 2008] Smet, I. D., Vassileva, V., Rybel, B. D., Levesque, M. P., Grunewald, W., Damme, D. V., Noorden, G. V., Naudts, M., Isterdael, G. V., Clercq, R. D., Wang, J. Y., Meuli, N., Vanneste, S., Friml, J., Hilson, P., JÃijrgens, G., Ingram, G. C., Inze, D., Benfey, P. N. and Beeckman, T. (2008). Receptor-like kinase ACR4 restricts formative cell divisions in the *Arabidopsis* root. Science *322*, 594–597.

[Solano et al., 1995] Solano, R., Nieto, C. and Paz-Ares, J. (1995). MYB.Ph3 transcription factor from *Petunia* hybrida induces similar DNA-bending/distortions on its two types of binding site. Plant J *8*, 673–682.

[Somerville and Browse, 1991] Somerville, C. and Browse, J. (1991). Plant lipids: metabolism, mutants, and membranes. Science *252*, 80–87.

[Staehelin, 1970] Staehelin, H. (1970). 4'-Demethyl-epipodophyllotoxin thenylidene glucoside (VM 26), a podophyllum compound with a new mechanism of action. Eur J Cancer *6*, 303–311.

[Staswick and Tiryaki, 2004] Staswick, P. E. and Tiryaki, I. (2004). The oxylipin signal jasmonic acid is activated by an enzyme that conjugates it to isoleucine in *Arabidopsis*. Plant Cell *16*, 2117–2127.

[Sterck et al., 2007] Sterck, L., Rombauts, S., Vandepoele, K., Rouze, P. and de Peer, Y. V. (2007). How many genes are there in plants and why are they there? Curr Opin Plant Biol *10*, 199–203.

[Stintzi et al., 1993] Stintzi, A., Heitz, T., Prasad, V., Wiedemann-Merdinoglu, S., Kauffmann, S., Geoffroy, P., Legrand, M. and Fritig, B. (1993). Plant 'pathogenesis-related' proteins and their role in defense against pathogens. Biochimie *75*, 687–706.

[Strong et al., 1997] Strong, S. J., Ohta, Y., Litman, G. W. and Amemiya, C. T. (1997). Marked improvement of PAC and BAC cloning is achieved using electroelution of pulsed-field gel-separated partial digests of genomic DNA. Nucleic Acids Res *25*, 3959–3961.

[Suffness, 1995] Suffness, M. (1995). Taxol: Science and Application. CRC Press, Inc., Boca Raton, FL.

[Sugimoto et al., 2010] Sugimoto, K., Jiao, Y. and Meyerowitz, E. M. (2010). *Arabidopsis* regeneration from multiple tissues occurs via a root development pathway. Dev Cell *18*, 463–471.

[Suzuki et al., 2002] Suzuki, G., Sook, G. and Mukai, Y. (2002). Efficient storage and screening system for onion BAC clones. Breeding Science *52*, 157–159.

[Suzuki et al., 2005] Suzuki, H., Reddy, M. S. S., Naoumkina, M., Aziz, N., May, G. D., Huhman, D. V., Sumner, L. W., Blount, J. W., Mendes, P. and Dixon, R. A. (2005). Methyl jasmonate and yeast elicitor induce differential transcriptional and metabolic re-programming in cell suspension cultures of the model legume *Medicago truncatula*. Planta *220*, 696–707.

['t Hoen et al., 2008] 't Hoen, P. A. C., Ariyurek, Y., Thygesen, H. H., Vreugdenhil, E., Vossen, R. H. A. M., de Menezes, R. X., Boer, J. M., van Ommen, G.-J. B. and den Dunnen, J. T. (2008). Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. Nucleic Acids Res *36*, e141.

[Tabata, 2004] Tabata, H. (2004). Paclitaxel production by Plant Cell Culture Technology, In: ADVANCES IN BIOCHEMICAL ENGINEERING/ BIOTECHNOLOGY. Springer Verlag, Berlin.

[Terauchi and Kahl, 2000] Terauchi, R. and Kahl, G. (2000). Rapid isolation of promoter sequences by TAIL-PCR: the 5-flanking regions of Pal and Pgi genes from yams. Mol Gen Genet *263*, 554–560.

[Theis and Lerdau, 2003] Theis, N. and Lerdau, M. (2003). The evolution of function in plant secondary metabolites. International Journal of Plant Sciences *164*, 93–102.

[Thines et al., 2007] Thines, B., Katsir, L., Melotto, M., Niu, Y., Mandaokar, A., Liu, G., Nomura, K., He, S. Y., Howe, G. A. and Browse, J. (2007). JAZ repressor proteins are targets of the SCF(COI1) complex during jasmonate signalling. Nature *448*, 661–665.

[Tsutsui et al., 2006] Tsutsui, T., Morita-Yamamuro, C., Asada, Y., Minami, E., Shibuya, N., Ikeda, A. and Yamaguchi, J. (2006). Salicylic acid and a chitin elicitor both control expression of the CAD1 gene involved in the plant immunity of *Arabidopsis*. Biosci Biotechnol Biochem *70*, 2042–2048.

[Ueda and Kato, 1980] Ueda, J. and Kato, J. (1980). Isolation and Identification of a Senescence-promoting Substance from Wormwood (*Artemisia absinthium L.*). Plant Physiol *66*, 246–249.

[van der Fits and Memelink, 2000] van der Fits, L. and Memelink, J. (2000). ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. Science *289*, 295–297.

[van der Fits and Memelink, 2001] van der Fits, L. and Memelink, J. (2001). The jasmonate-inducible AP2/ERF-domain transcription factor ORCA3 activates gene expression via interaction with a jasmonate-responsive promoter element. Plant J *25*, 43–53.

[van der Graaff et al., 2000] van der Graaff, E., Dulk-Ras, A. D., Hooykaas, P. J. and Keller, B. (2000). Activation tagging of the *LEAFY PETIOLE* gene affects leaf petiole development in *Arabidopsis thaliana*. Development *127*, 4971–4980.

[van der Heijden et al., 2004] van der Heijden, R., Jacobs, D. I., Snoeijer, W., Hallard, D. and Verpoorte, R. (2004). The Catharanthus alkaloids: pharmacognosy and biotechnology. Curr Med Chem *11*, 607–628.

[van Rozendaal et al., 2000] van Rozendaal, E. L., Lelyveld, G. P. and van Beek, T. A. (2000). Screening of the needles of different yew species and cultivars for paclitaxel and related taxoids. Phytochemistry *53*, 383–389.

[VanAlstyne et al., 1999] VanAlstyne, K. L., III, J. J. M., Hustead, C. L. and Kearns, L. J. (1999). Phlorotannin allocation among tissues of northeast pacific Kelps and Rockweeds. J. Phycol. *35*, 483–492.

[Vera et al., 2008] Vera, J. C., Wheat, C. W., Fescemyer, H. W., Frilander, M. J., Crawford, D. L., Hanski, I. and Marden, J. H. (2008). Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. Mol Ecol *17*, 1636–1647.

[Vidensek et al., 1990] Vidensek, N., Lim, P., Campbell, A. and Carlson, C. (1990). Taxol content in bark, wood, root, leaf, twig, and seedling from several *Taxus* species. J Nat Prod *53*, 1609–1610.

[Voelkerding et al., 2009] Voelkerding, K. V., Dames, S. A. and Durtschi, J. D. (2009). Next-generation sequencing: from basic research to diagnostics. Clin Chem *55*, 641–658.

[Voinnet et al., 2003] Voinnet, O., Rivas, S., Mestre, P. and Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J *33*, 949–956.

[von Siebold and Zuccarini, 1870] von Siebold, P. F. and Zuccarini, J. G. (1870). Flora Japonica.

[Vongpaseuth et al., 2007] Vongpaseuth, K., Nims, E., Amand, M. S., Walker, E. L. and Roberts, S. C. (2007). Development of a particle bombardment-mediated transient transformation system for *Taxus spp.* cells in culture. Biotechnol Prog *23*, 1180–1185.

[Walhout et al., 2000] Walhout, A. J. M., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S. and Vidal, M. (2000). Application to the cloning of large numbers of open reading frames or ORFeomes. Methods Enzymol. *328*, 575–592.

[Walker and Croteau, 2000a] Walker, K. and Croteau, R. (2000a). Taxol biosynthesis: molecular cloning of a benzoyl-CoA:taxane 2alpha-O-benzoyltransferase cDNA from *Taxus* and functional expression in *Escherichia coli*. Proc Natl Acad Sci U S A *97*, 13591–13596.

[Walker and Croteau, 2000b] Walker, K. and Croteau, R. (2000b). Molecular cloning of a 10-deacetylbaccatin III-10-O-acetyl transferase cDNA from *Taxus* and functional expression in *Escherichia coli*. Proc Natl Acad Sci U S A *97*, 583–587.

[Walker and Croteau, 2001] Walker, K. and Croteau, R. (2001). Taxol biosynthetic genes. Phytochemistry *58*, 1–7.

[Walker et al., 2002] Walker, K., Fujisaki, S., Long, R. and Croteau, R. (2002). Molecular cloning and heterologous expression of the C-13 phenylpropanoid

side chain-CoA acyltransferase that functions in Taxol biosynthesis. Proc Natl Acad Sci U S A *99*, 12715–12720.

[Walker et al., 1999] Walker, K., Ketchum, R. E., Hezari, M., Gatfield, D., Goleniowski, M., Barthol, A. and Croteau, R. (1999). Partial purification and characterization of acetyl coenzyme A: taxa-4(20),11(12)-dien-5alpha-ol O-acetyl transferase that catalyzes the first acylation step of taxol biosynthesis. Arch Biochem Biophys *364*, 273–279.

[Walker et al., 2002] Walker, K., Long, R. and Croteau, R. (2002). The final acylation step in taxol biosynthesis: cloning of the taxoid C13-side-chain N-benzoyltransferase from *Taxus*. Proc Natl Acad Sci U S A *99*, 9166–9171.

[Wall et al., 1966] Wall, M., Wani, M., Cook, C. and Palmer, K. (1966). Plant antitumor agents. I. The isolation and structure of camptothecin, a novel alkaloidal leukemia and tumor inhibitor from *Camptotheca acuminata*. Journal of the American Chemical Society *88*, 3888–3890.

[Wan and Wilkins, 1993] Wan, C. Y. and Wilkins, T. A. (1993). Spermidine facilitates PCR amplification of target DNA. PCR Methods Appl *3*, 208–210.

[Wang et al., 2001] Wang, C., Wu, J. and Mei, X. (2001). Enhanced taxol production and release in *Taxus chinensis* cell suspension cultures with selected organic solvents and sucrose feeding. Biotechnol Prog *17*, 89–94.

[Wang et al., 2009] Wang, H., Zhu, Y., Fujioka, S., Asami, T., Li, J. and Li, J. (2009). Regulation of *Arabidopsis* brassinosteroid signaling by atypical basic helix-loop-helix proteins. Plant Cell *21*, 3781–3791.

[Wang et al., 2010] Wang, Y. F., Yu, S. H., Dong, M., Zhang, M. L., Huo, C. H. and Shi, Q. W. (2010). Chemical Studies on *Taxus cuspidata*. Chem Biodivers *7*, 1698–1716.

[Wani et al., 1971] Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P. and McPhail, A. T. (1971). Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. J Am Chem Soc *93*, 2325–2327.

[Wasternack, 2007] Wasternack, C. (2007). Jasmonates: an update on biosynthesis, signal transduction and action in plant stress response, growth and development. Ann Bot *100*, 681–697.

[Weake and Workman, 2010] Weake, V. M. and Workman, J. L. (2010). Inducible gene expression: diverse regulatory mechanisms. Nat Rev Genet *11*, 426–437.

[Weiler et al., 1998] Weiler, E., Laudert, D., Schaller, F., Stelmach, B. and Hennig, P. (1998). Fatty acid-derived signaling molecules in the interaction of plants with their environment. RECENT ADVANCES IN PHYTOCHEMISTRY *32*, 179–206.

[Weltmeier et al., 2009] Weltmeier, F., Rahmani, F., Ehlert, A., Dietrich, K., Schuetze, K., Wang, X., Chaban, C., Hanson, J., Teige, M., Harter, K., Vicente-Carbajosa, J., Smeekens, S. and DrÃűge-Laser, W. (2009). Expression patterns within the *Arabidopsis* C/S1 bZIP transcription factor network: availability of heterodimerization partners controls gene expression during stress response and development. Plant Mol Biol *69*, 107–119.

[Wheeler et al., 2008] Wheeler, D. A., Srinivasan, M., Egholm, M., Shen, Y., Chen, L., McGuire, A., He, W., Chen, Y. J., Makhijani, V., Roth, G. T., Gomes, X., Tartaro, K., Niazi, F., Turcotte, C. L., Irzyk, G. P., Lupski, J. R., Chinault, C., Song, X., Liu, Y., Yuan, Y., Nazareth, L., Qin, X., Muzny, D. M., Margulies, M., Weinstock, G. M., Gibbs, R. A. and Rothberg, J. M. (2008). The complete genome of an individual by massively parallel DNA sequencing. Nature *452*, 872–876.

[Wildung and Croteau, 1996] Wildung, M. R. and Croteau, R. (1996). A cDNA clone for taxadiene synthase, the diterpene cyclase that catalyzes the committed step of taxol biosynthesis. J Biol Chem *271*, 9201–9204.

[Williams et al., 2000] Williams, D. C., Wildung, M. R., Jin, A. Q., Dalal, D., Oliver, J. S., Coates, R. M. and Croteau, R. (2000). Heterologous expression and characterization of a "Pseudomature" form of taxadiene synthase involved in paclitaxel (Taxol) biosynthesis and evaluation of a potential intermediate and inhibitors of the multistep diterpene cyclization reaction. Arch Biochem Biophys *379*, 137–146.

[Wink, 1985] Wink, M. (1985). Plant breeding: importance of plant secondary metabolites for protection against pathogens and herbivores. P1. Syst. Evol. *150*, 65–81.

[Witherup et al., 1990] Witherup, K. M., Look, S. A., Stasko, M. W., Ghiorzi, T. J., Muschik, G. M. and Cragg, G. M. (1990). Taxus spp. needles contain amounts of taxol comparable to the bark of *Taxus brevifolia*: analysis and isolation. J Nat Prod *53*, 1249–1255.

[Wu and Ge, 2004] Wu, J. and Ge, X. (2004). Oxidative burst, jasmonic acid biosynthesis, and taxol production induced by low-energy ultrasound in Taxus chinensis cell suspension cultures. Biotechnol Bioeng *85*, 714–721.

[Wu and Lin, 2003] Wu, J. and Lin, L. (2003). Enhancement of taxol production and release in *Taxus chinensis* cell cultures by ultrasound, methyl jasmonate and in situ solvent extraction. Appl Microbiol Biotechnol *62*, 151–155.

[Xie et al., 1998] Xie, D. X., Feys, B. F., James, S., Nieto-Rostro, M. and Turner, J. G. (1998). *COI1*: an *Arabidopsis* gene required for jasmonate-regulated defense and fertility. Science *280*, 1091–1094.

[Xu and Timko, 2004] Xu, B. and Timko, M. (2004). Methyl jasmonate induced expression of the tobacco putrescine N -methyltransferase genes requires both G-box and GCC-motif elements. Plant Mol Biol *55*, 743–761.

[Xu et al., 1998] Xu, J. F., Yin, P. Q., Wei, X. G. and Su, Z. G. (1998). Self-immobilized aggregate culture of *Taxus cuspidata* for improved taxol production. Biotechnology Techniques *12*, 241–244.

[Xu et al., 2002] Xu, L., Liu, F., Lechner, E., Genschik, P., Crosby, W. L., Ma, H., Peng, W., Huang, D. and Xie, D. (2002). The SCF(COI1) ubiquitin-ligase complexes are required for jasmonate response in *Arabidopsis*. Plant Cell *14*, 1919–1935.

[Xu et al., 2004] Xu, M., Dong, J. and Zhu, M. (2004). Involvement of NO in fungal elicitor-induced activation of PAL and stimulation of taxol synthesis

in *Taxus chinensis* suspension cells. Biomedical and Life Sciences *49*, 1038–1043.

[Yamaguchi-Shinozaki and Shinozaki, 2005] Yamaguchi-Shinozaki, K. and Shinozaki, K. (2005). Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. Trends Plant Sci *10*, 88–94.

[Yamamoto et al., 2001] Yamamoto, R., Fujioka, S., Demura, T., Takatsuto, S., Yoshida, S. and Fukuda, H. (2001). Brassinosteroid levels increase drastically prior to morphogenesis of tracheary elements. Plant Physiol *125*, 556–563.

[Yan et al., 2007] Yan, Y., Stolz, S., Chetelat, A., Reymond, P., Pagni, M., Dubugnon, L. and Farmer, E. E. (2007). A downstream mediator in the growth repression limb of the jasmonate pathway. Plant Cell *19*, 2470–2483.

[Yang et al., 2010] Yang, P., Li, X., Shipp, M. J., Shockey, J. M. and Cahoon, E. B. (2010). Mining the bitter melon (*Momordica charantia l.*) seed transcriptome by 454 analysis of non-normalized and normalized cDNA populations for conjugated fatty acid metabolism-related genes. BMC Plant Biol *10*, 250.

[Yang et al., 2000] Yang, Y., Li, R. and Qi, M. (2000). In vivo analysis of plant promoters and transcription factors by agroinfiltration of tobacco leaves. Plant J *22*, 543–551.

[Yeung et al., 1999] Yeung, T. K., Germond, C., Chen, X. and Wang, Z. (1999). The mode of action of taxol: apoptosis at low concentration and necrosis at high concentration. Biochem Biophys Res Commun *263*, 398–404.

[Yu et al., 2000] Yu, Y., Tomkins, J., Waugh, R., Frisch, D., Kudrna, D., Kleinhofs, A. and Wing, R. (2000). A bacterial artificial chromosome library from barley (*Hordeum vulgare*). Theor Appl Genet *101*, 1093–1099.

[Zhang et al., 2000] Zhang, C., Mei, X., Liu, L. and Yu, L. (2000). Enhanced paclitaxel production induced by the combination of elicitors in cell suspension cultures of *Taxus chinensis*. Biotechnology Letters *22*, 1561–1564.

[Zhao and Last, 1996] Zhao, J. and Last, R. L. (1996). Coordinate regulation of the tryptophan biosynthetic pathway and indolic phytoalexin accumulation in *Arabidopsis*. Plant Cell *8*, 2235–2244.

[Zhou et al., 1997] Zhou, J., Tang, X. and Martin, G. B. (1997). The Pto kinase conferring resistance to tomato bacterial speck disease interacts with proteins that bind a cis-element of pathogenesis-related genes. EMBO J *16*, 3207–3218.

[Zhulidov et al., 2004] Zhulidov, P. A., Bogdanova, E. A., Shcheglov, A. S., Vagner, L. L., Khaspekov, G. L., Kozhemyako, V. B., Matz, M. V., Meleshkevitch, E., Moroz, L. L., Lukyanov, S. A. and Shagin, D. A. (2004). Simple cDNA normalization using Kamchatka crab duplex-specific nuclease. Nucleic Acids Res *32*, e37.

[Zulak et al., 2007] Zulak, K. G., Cornish, A., Daskalchuk, T. E., Deyholos, M. K., Goodenowe, D. B., Gordon, P. M. K., Klassen, D., Pelcher, L. E., Sensen, C. W. and Facchini, P. J. (2007). Gene transcript and metabolite profiling of elicitor-induced opium poppy cell cultures reveals the coordinate regulation of primary and secondary metabolism. Planta *225*, 1085–1106.

[Zuraida et al., 2010] Zuraida, A. R., Rahiniza, K., Hafiza, N., Roowi, S., Zamri, Z. and Subramaniam, S. (2010). Factors affecting delivery and transient expression of gusA gene in Malaysian indica rice MR 219 callus via biolistic gun system. African Journal of Biotechnology *9*, 8810–8818.

Appendix I

TED2 sequence alignment

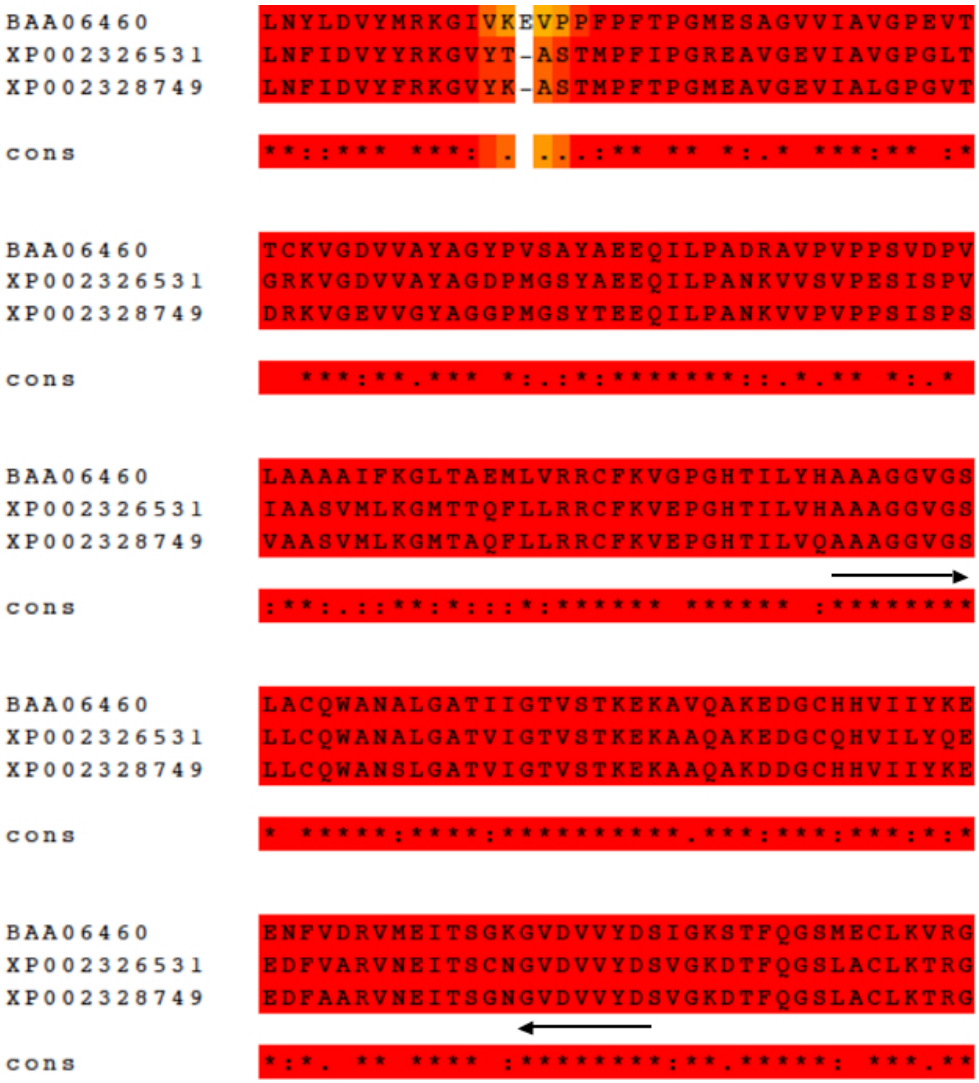

Figure 6.4: Relevant part of the multiple protein sequence alignment of the authentic TED2 from *Zinnea elegans* [Accesion: BAA06460] and two homologues from *Populus trichocarpa* [Accession: XP002326531 and XP002328749] used for degenerate primer design. Sequences used for the forward and reverse primer design are indicated.

Mixed base definitions for degenerate primer design

Code letter	Base
R	A,G
Y	C, T
M	A, C
K	G, T
S	C, G
W	A, T
H	A, C, T
B	C, G, T
V	A, C, G
D	A, G, T
N	A, C, G, T

Table 6.1: The letter codes were used for the design of degenerate primers to amplify the *TED2* orthologue from *T. cuspidata* genomic DNA

Appendix II

Probability of BAC library screening

The probability to find a target gene within a constructed genomic library can be calculated with the following equation:

$\boldsymbol{P} = \mathbf{1\text{-}e}^{\boldsymbol{N}[ln(1-I/GS)]}$

N represents the number of BAC clones required = X

I represents the average insert size = 73 kb

GS is the haploid genome size = 10850 Mb

This formula was applied to calculate the number of required cfus harbouring a BAC construct with a probability of 98% to find the gene of interest within the genomic library.

[1] It is necessary to rearrange the formula to calculate the number of cfu where N is equal to the natural log of 1-P.

N[ln(1-I/GS)] = ln(1-P)

[2] To isolate N the formula is further rearranged

N = (ln(1-P) / (ln 1-I/GS)

N = (ln(1-0.98) / (ln 1-73 kb/10,850,000)

N = ln 0.02 / ln 1-6.728 x 10^{-6}

N = ln 0.02 / ln 0.9999932719

N = ln 0.02 / -6.728133233 x 10^{-6}

N = -3.912023005 / -6.728133233 x 10^{-6}

$\boldsymbol{N = 581{,}422}$ **cfu**

Applying this calculation, the required number of cfu of which all harbour a BAC construct with an averge insert size of 73 kb of *T. cuspidata* genomic DNA is 581,422 to have a 98% probability to find the gene of interest within the library.

Appendix III

A. thaliana homologues of d.e. contigs in the procambium csc

Contig	Annotation	Plant species	log2 FC	*A. thaliana* hom.	E-value	Protein interaction
25515	Myb TF	*P. sitchensis*	3.3	AT5G47390	5×10^{-5}	no
22973	ERF TF	*P. salicina*	3.8	n/a	n/a	n/a
10106	bZIP TF	*G. max*	2.3	AT2G16770	4×10^{-57}	no
18201	Myb TF	*H. brasiliensiis*	1.8	n/a	n/a	n/a
13935	putative TF	*R. communis*	-2.3	AT3G17100	9×10^{-5}	AtBS1
33050	Myb TF	*G. hybrida*	-3.2	n/a	n/a	n/a
15401	Myb TF	*O. sativa*	-3.8	AAM67076.1	9×10^{-38}	no
13174	AP2/ERF TF	*P. trichocarpa*	-6.8	AT5G13330	3×10^{-26}	no

Appendix IV

Shared d.e. contigs with a JA signalling associated GO term

		procambium csc [p.e.]			somatic csc [p.e.]					
Contig	GO term	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.	Annotation	*A. thaliana* hom.	E-value
14280	response to JA stim.	-0.9	-0.3	-2.6	1.1	0.1	-0.1	JAZ10	AT5G13220	2×10^{-8}
11289	JA mediated sig. path.	1.87	3.7	3.43	2.2	3.2	2.4	JAZ3	AT3G17860	8×10^{-13}
33381	JA mediated sig. path.	0.6	2.8	2.7	0.8	3.8	3.6	JAZ2	At1G74950	7×10^{-9}
22910	JA mediated sig. path.	0.6	3.2	1.6	1.2	2.7	3.7	*Nt*JAZ3	n/a	9×10^{-17}

Appendix V: *T. cuspidata* TFs

up-regulated AP2/ERF TFs

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
significant log2 fold-change at 0.5 h p.e.			
02463	**6.1**	**-3.0**	**1.6**
00499	**4.9**	-0.8	-0.8
22386	**4.5**	**1.0**	**-5.8**
04485	**4.4**	0.3	0.3
14838	**3.1**	-0.7	-2.2
26170	**2.3**	-0.8	**-3.6**
07245	**2.3**	0.3	**-3.6**
16850	**2.0**	0.1	**-1.6**
03304	**1.7**	-0.6	0.1
22558	**1.5**	0.6	**-1.7**
17139	**1.5**	-0.9	-0.4
10213	**1.5**	0.3	-0.3
01431	**1.5**	**-1.1**	0.6
25770	**1.2**	2.8	**-1.2**
significant log2 fold-change at 2 h p.e.			
13174	0.2	**2.9**	0.1
00242	0.7	**1.3**	**-1.7**
significant log2 fold-change at 12 h p.e.			
03038	-0.2	0.6	**3.5**
00241	0.6	-0.9	**2.1**
17529	0.8	-0.8	**1.6**

down-regulated AP2/ERF TFs

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
significant log2 fold-change at 0.5 h p.e.			
31207	**-2.1**	0.0	**2.3**
22973	**-5.2**	**1.3**	**1.5**
17789	**-3.6**	0.6	**-2.9**
17303	**-1.6**	0.2	-0.1
17212	**-3.2**	-0.2	**1.3**
13937	**-1.4**	**1.3**	**-1.5**

down-regulated AP2/ERF TFs

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
significant log2 fold-change at 2 h p.e.			
04977	-0.1	**-1.1**	**1.5**
significant log2 fold-change at 12 h p.e.			
18881	-0.1	-0.7	**-6.3**
18880	0.0	0.2	**-1.0**
26842	-0.2	0.8	**-1.0**

AP2/ERF TFs with no significant fold change

34146	-0.5	-0.2	-0.3
32800	-0.4	0.2	-0.6
27511	-0.1	0.1	-0.5
26196	0.0	0.3	-0.3
26194	0.0	0.1	-0.1
26133	0.0	0.0	0.0
25513	0.0	0.6	-0.6
24081	0.1	0.0	0.6
22332	-0.4	0.5	-0.4
21469	0.8	0.3	-0.8
21348	0.0	0.1	-0.1
18166	0.1	0.6	-0.7
17753	-0.2	-0.2	-0.4
16735	-0.3	0.3	-0.4
14538	0.4	0.5	0.4
13601	-0.5	0.8	0.4
10324	-0.1	-0.8	-1.4
09700	0.4	-0.4	-0.8
08181	-0.8	-0.5	-0.7
08057	0.0	0.0	-0.1
03962	-0.9	0.7	-1.7
02462	0.0	0.2	0.0
02459	0.7	-0.6	0.0
02285	0.4	-0.5	-0.2
01900	-0.2	0.4	0.8
01858	-0.4	-0.2	0.1
00114	0.0	0.2	-0.2

up-regulated bHLH TFs

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
significant log2 fold-change at 0.5 h p.e.			
11748	**2.4**	0.2	**-1.2**
02991	**2.3**	**2.3**	**-1.2**
13935	**2.1**	0.7	-0.1
08058	**1.5**	0.2	0.1
15303	**1.0**	0.7	-0.5
significant log2 fold-change at 2 h p.e.			
02611	-0.6	**2.8**	-0.3
19503	0.6	**2.0**	**-2.2**
12353	-0.8	**1.4**	**-1.2**
03153	0.2	**1.4**	**-1.8**
15240	0.5	**1.1**	**-2.5**
significant log2 fold-change at 12 h p.e.			
00987	0.7	0.9	**1.1**

down-regulated bHLH TFs

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
significant log2 fold-change at 0.5 h p.e.			
22017	**-2.8**	0.4	-3.6
19897	**-2.5**	-0.2	**-2.4**
31235	**-2.3**	**1.3**	**3.4**
01822	**-2.3**	0.4	-0.9
17793	**-1.9**	-0.4	0.9
20806	**-1.3**	-0.1	-0.2
20044	**-1.7**	-0.1	**-2.2**
02439	**-1.4**	-0.2	**-1.8**
16264	**-1.1**	0.6	-0.6
significant log2 fold-change at 2 h p.e.			
17494	0.2	**-1.6**	-0.4
19154	-0.1	**-1.1**	1.1

down-regulated bHLH TFs

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
significant log2 fold-change at 12 h p.e.			
10889	0.9	0.6	**-2.6**
12598	-0.8	0.3	**-2.3**
19109	0.3	-0.8	**-2.1**
04335	-0.8	-0.7	**-1.9**
21972	0.1	-0.2	**-1.8**
02498	-0.5	0.6	**-1.7**
02518	-0.1	0.6	**-1.5**
32372	0.3	0.3	**-1.3**
22095	0.0	1.0	**-1.3**
14220	-0.2	-0.8	**-1.1**
15365	0.3	0.7	**-1.0**

bHLH TFs with no significant fold-change

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
33006	-0.1	-0.2	-0.9
28045	0.1	0.2	-0.4
21671	-0.1	0.4	0.2
19723	-0.3	1.0	-0.1
12354	-0.4	0.4	-0.3
16478	0.2	0.6	0.3
05557	-0.5	-1.0	-0.3
02514	0.2	0.0	-0.2
02513	0.0	-0.1	0.0
02512	0.0	0.6	-0.6
01083	-0.8	0.7	-0.7

up-regulated WRKY TFs

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
significant log2 fold-change at 0.5 h p.e.			
19284	**3.6**	0.5	**1.3**
09595	**2.3**	0.2	**1.8**
27015	**1.6**	**-1.2**	**-1.1**
10530	**1.5**	**-1.4**	0.9
26745	**1.1**	**-1.0**	-0.5
23796	**1.1**	**1.2**	0.0
09027	**1.0**	0.1	-0.8
significant log2 fold-change at 2 h p.e.			
04263	0.8	**3.0**	0.8
20584	0.4	**1.0**	**-1.0**

down-regulated WRKY TFs

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
significant log2 fold-change at 0.5 h p.e.			
06574	**-1.2**	0.6	**-1.4**
19093	**-1.1**	-0.2	**1.0**
significant log2 fold-change at 2 h p.e.			
21172	-0.1	**-1.2**	-0.5
significant log2 fold-change at 12 h p.e.			
04857	-0.4	0.5	**-2.2**
16617	-0.4	-0.4	**-1.8**
10839	-0.5	0.7	**-1.4**
15956	-0.5	-0.8	**-1.3**
29087	0.5	0.7	**-1.2**

WRKY TFs with no significant fold-change

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
35166	0.0	0.2	-0.1
32598	0.0	0.2	-0.2
32419	-0.3	-0.4	-0.2
31780	0.0	0.0	0.0
29217	-0.8	0.1	-0.3
28143	0.0	0.0	0.0
24944	0.3	-0.1	-0.4
24699	0.0	0.0	0.0
24463	-0.1	-0.3	-1.0
23553	-0.1	0.1	-0.1
21666	-0.1	0.5	-1.0
21078	0.5	-0.2	-0.4
21048	-0.4	-0.1	0.2
19509	-0.1	0.0	-0.6
19392	-0.1	-0.2	-0.5
19342	-0.1	-0.3	0.2
19260	-0.9	-0.4	0.3
14650	0.2	-0.2	-0.3
14506	0.1	0.1	0.1
12889	-0.1	0.2	-0.6
11929	-0.8	0.0	0.0
11928	0.0	0.1	-0.1
11676	-0.3	0.0	0.0
09626	0.3	-0.2	-0.5
07935	-0.1	-0.1	-0.8
02885	0.3	-0.3	-0.7

up-regulated MYB TFs

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
significant log2 fold-change at 0.5 h p.e.			
05913	**3.3**	**-1.6**	-1.1
21458	**3.1**	-0.7	0.4
12379	**3.0**	0.0	**-2.4**
10258	**3.0**	-0.8	**-1.9**
10855	**2.7**	-0.3	0.0
10385	**2.6**	-0.9	-0.3
15401	**2.4**	0.5	0.9
12425	**2.1**	**2.0**	**-3.2**
06799	**1.8**	-0.4	-0.2
16463	**1.7**	-0.7	0.5
16408	**1.3**	-0.8	**-1.5**
20421	**1.3**	**1.7**	**-2.2**
14612	**1.3**	-0.2	**-1.9**
33050	**1.2**	0.3	**-2.1**
23804	**1.1**	0.3	**-1.3**
significant log2 fold-change at 2 h p.e.			
27845	0.4	**1.1**	**-1.3**

down-regulated MYB TFs

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
significant log2 fold-change at 0.5 h p.e.			
18201	**-1.4**	**-1.3**	0.3
significant log2 fold-change at 2 h p.e.			
26508	0.8	**-1.9**	**1.0**
06305	-0.1	**-1.2**	**-1.9**
significant log2 fold-change at 12 h p.e.			
09261	0.6	-0.8	**-1.8**
26217	0.3	0.7	**-1.5**
23200	0.6	-0.1	**-1.3**

MYB TFs with no significant fold-change

Contig	log2 FC 0.4h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
33324	0.0	0.2	-0.2
29875	0.0	0.2	-0.2
28128	-0.9	-0.4	-0.3
28055	0.0	0.5	-0.8
27559	0.3	0.2	-0.4
26516	0.1	-0.5	0.2
26386	0.8	0.9	-0.4
25515	-0.5	-1.0	-0.6
24985	0.1	0.3	0.4
23296	0.0	0.0	0.0
22158	-0.5	0.0	-0.8
21477	0.1	0.1	-0.2
25515	-0.5	-1.0	-0.6
24985	0.1	0.3	0.4
23296	0.0	0.0	0.0
22158	-0.5	0.0	-0.8
21477	0.1	0.1	-0.2
20533	-0.1	0.1	-0.1
17959	0.3	-0.5	-0.4
17907	0.8	-0.5	0.1
15537	-0.3	0.0	-0.8
13883	0.0	0.2	-0.2
11102	0.5	0.0	-0.9
02895	0.9	-0.6	0.1

up-regulated bZIP TFs

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
significant log2 fold-change at 0.5 h p.e.			
02478	**1.4**	-0.6	**-4.5**
12936	**1.2**	-0.5	**-3.3**
significant log2 fold-change at 2 h p.e.			
00580	-0.6	**1.5**	-0.5
06766	0.8	**1.3**	**-2.3**
significant log2 fold-change at 12 h p.e.			
18322	-0.1	-0.4	**1.3**

down-regulated bZIP TFs

Contig	log2 FC 0.5h	log2 FC 2h p.e.	log2 FC 12h p.e.
significant log2 fold-change at 0.5 h p.e.			
29275	**-3.3**	0.2	**-3.0**
05234	**-1.3**	0.0	**1.1**
significant log2 fold-change at 2 h p.e.			
12221	-0.3	**-1.3**	-0.2
10106	0.1	**-1.5**	0.3
24743	-0.2	**-1.1**	**-1.0**
significant log2 fold-change at 12 h p.e.			
22060	-0.6	-0.1	**-1.5**
14809	-0.5	0.6	**-1.7**
08495	-0.7	0.5	**-1.1**
02143	-0.9	0.6	**-1.6**

bZIP TFs with no significant fold-change

Contig	log2 FC 0.5h p.e.	log2 FC 2h p.e.	log2 FC 12h p.e.
28674	0.2	-0.2	0.0
22591	-0.6	-0.5	-0.1
17618	-0.1	0.3	-0.3
16097	-0.7	0.3	-0.2
15895	-0.9	0.4	-0.5
15750	-0.4	0.2	-0.2
11083	-0.6	0.0	0.1
10140	0.3	-0.1	-0.9
08012	-0.3	0.6	-0.5
06216	-0.9	0.9	-0.5
05322	-0.3	-0.2	0.0
02039	0.0	-0.3	-0.4
00761	-0.7	-0.2	0.6

Appendix VI: AP2/ERF TFs expression profiles

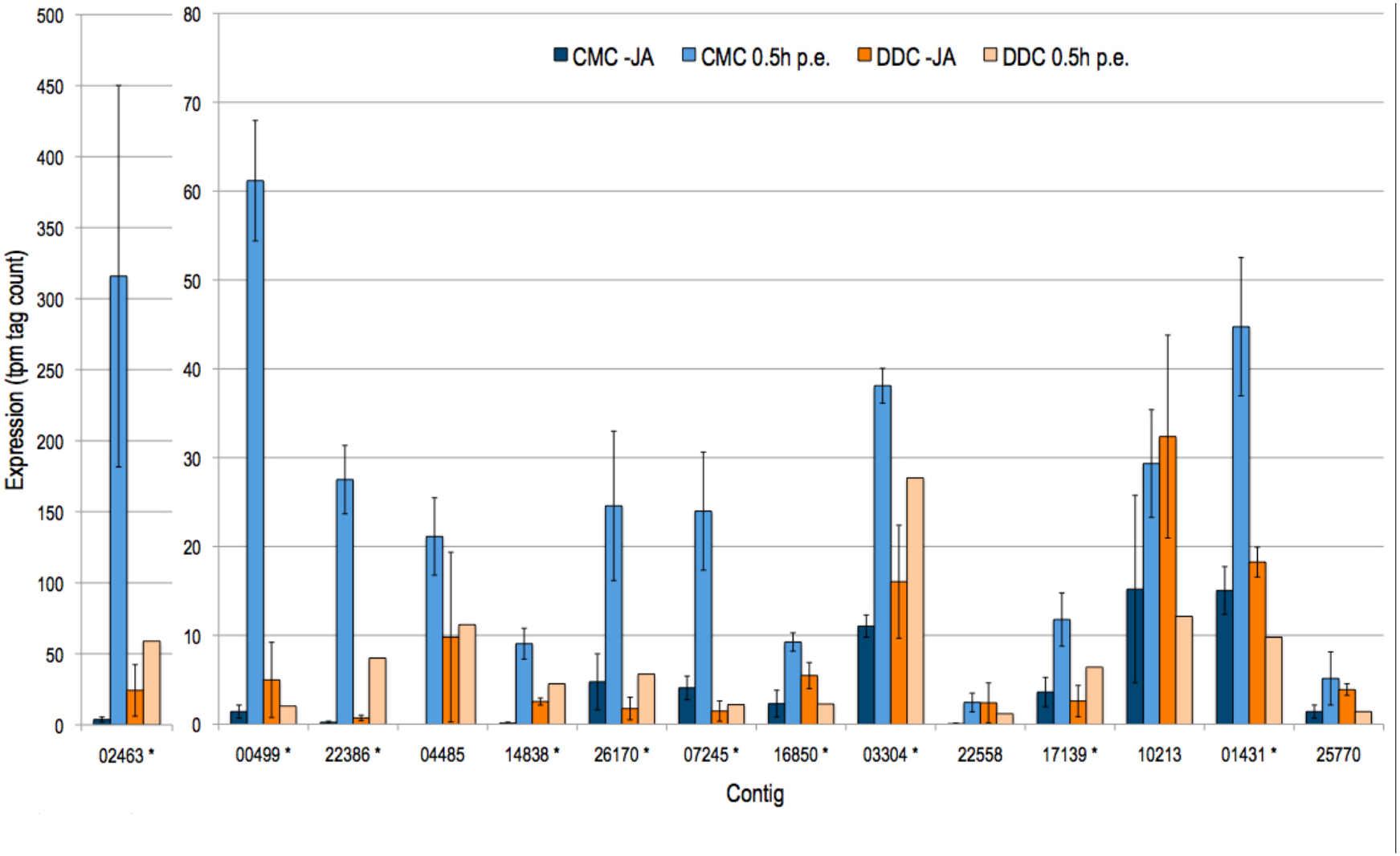

Figure 6.5: Expression levels of identified AP2/ERF TFs in CMCs and DDCs which show a significant fold change at **0.5 hours p.e.**. Asterisks indicates contigs with a differential higher expression profile in CMCs.

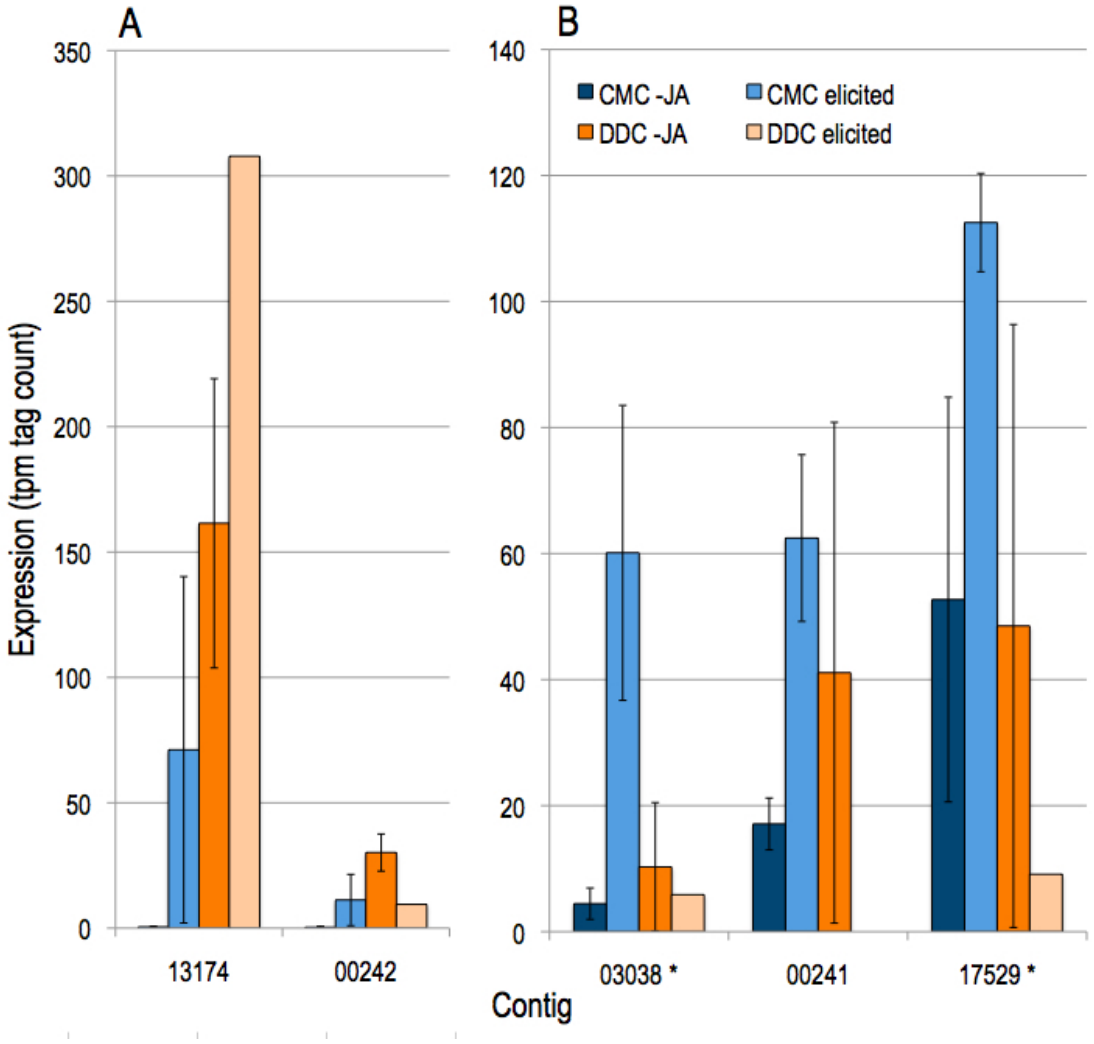

Figure 6.6: Expression levels of identified AP2/ERF TFs in CMCs and DDCs which show a significant fold change at [A]: **2 hours p.e.** and [B]: **12 hours p.e.** Asterisks indicates contigs with a differential higher expression profile in CMCs.

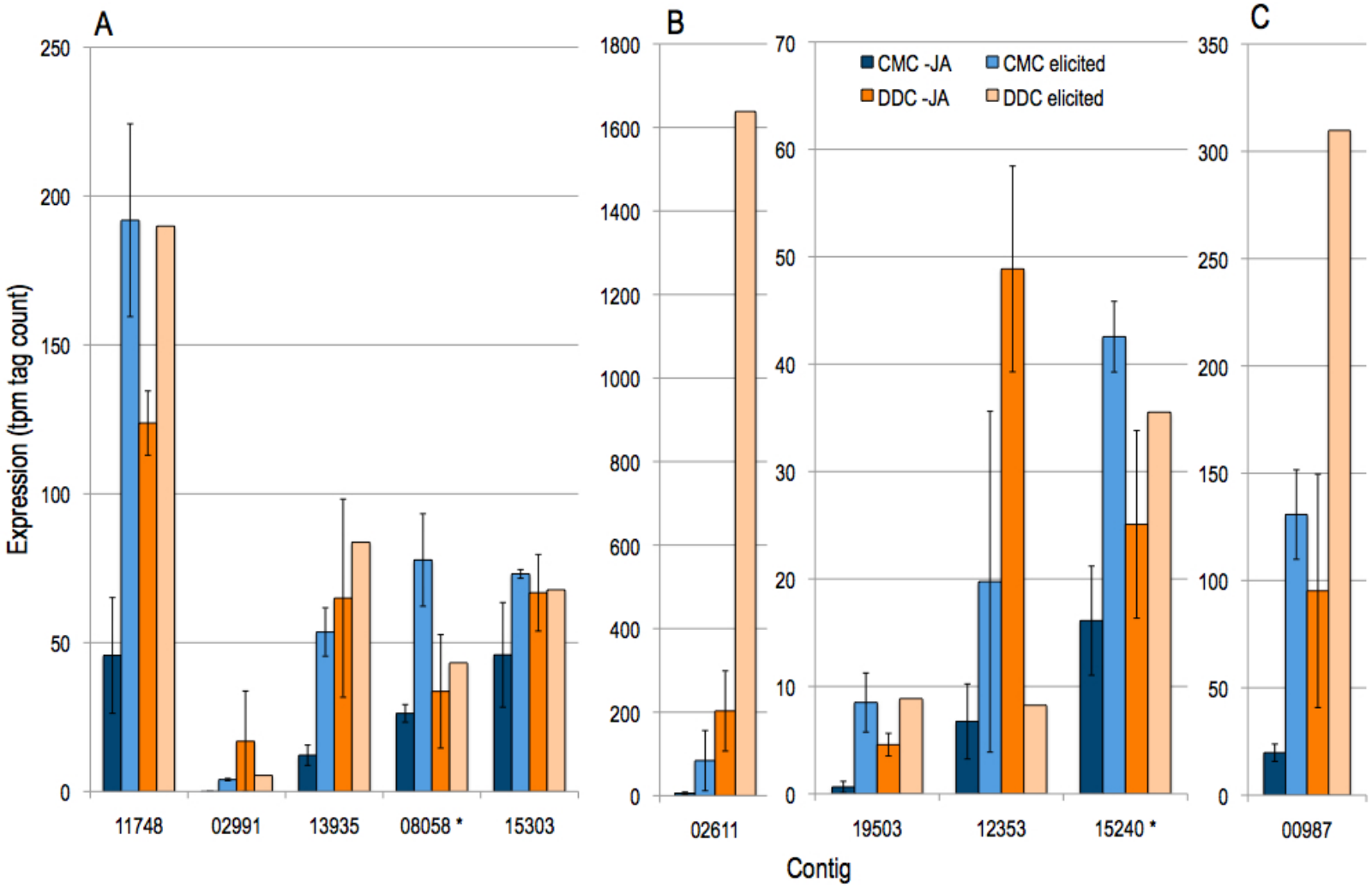

Figure 6.7: Expression levels of identified bHLH TFs in CMCs and DDCs which show a significant fold change at [A]: **0.5 hours p.e.**, [B]: **2 hours p.e.** and [C]: **12 hours p.e.** Asterisks indicates contigs with a differential higher expression profile in CMCs.

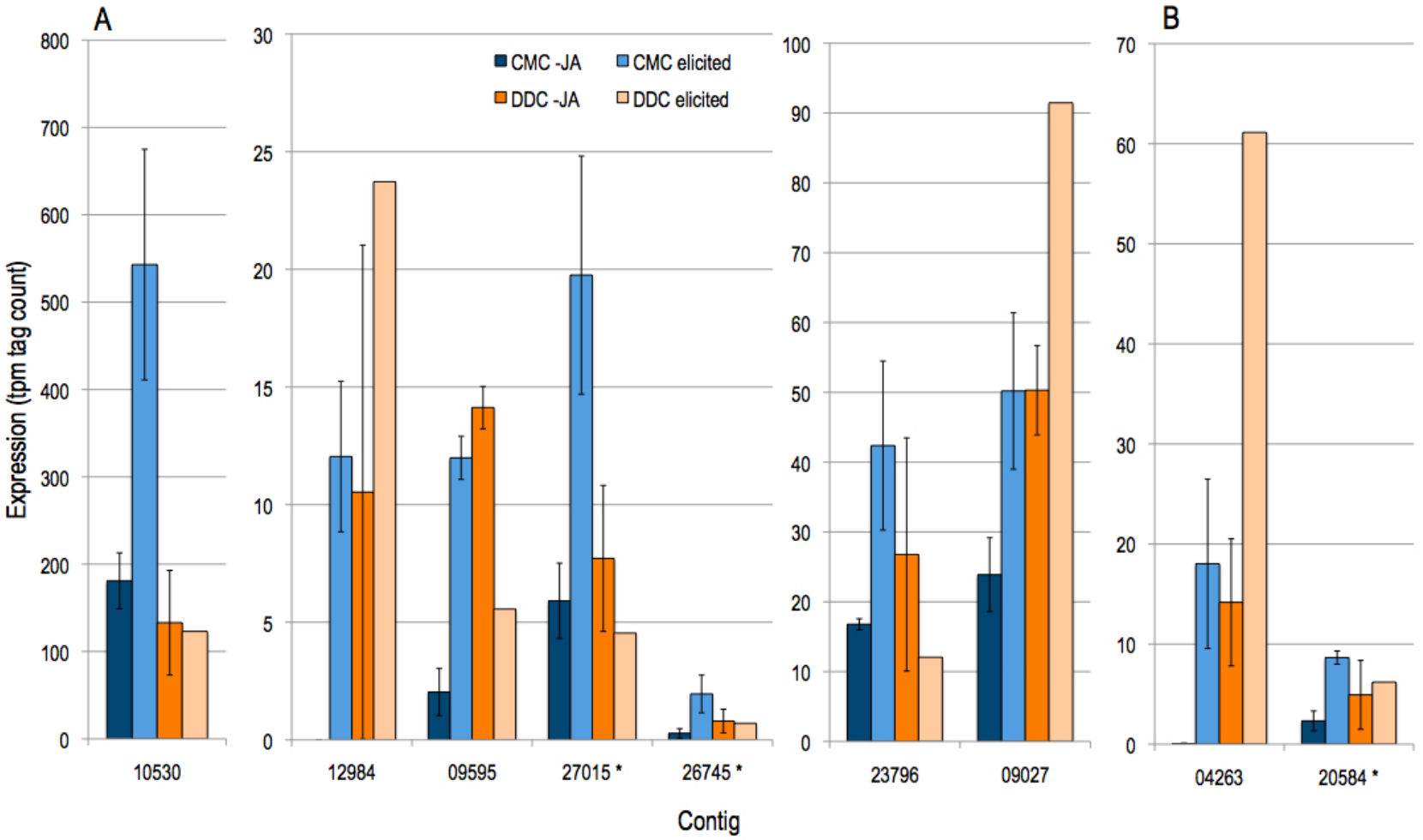

Figure 6.8: Expression levels of identified WRKY TFs in CMCs and DDCs which show a significant fold change at [A]: **0.5 hours p.e.** and [B]: **2 hours p.e.**. Asterisks indicates contigs with a differential higher expression profile in CMCs.

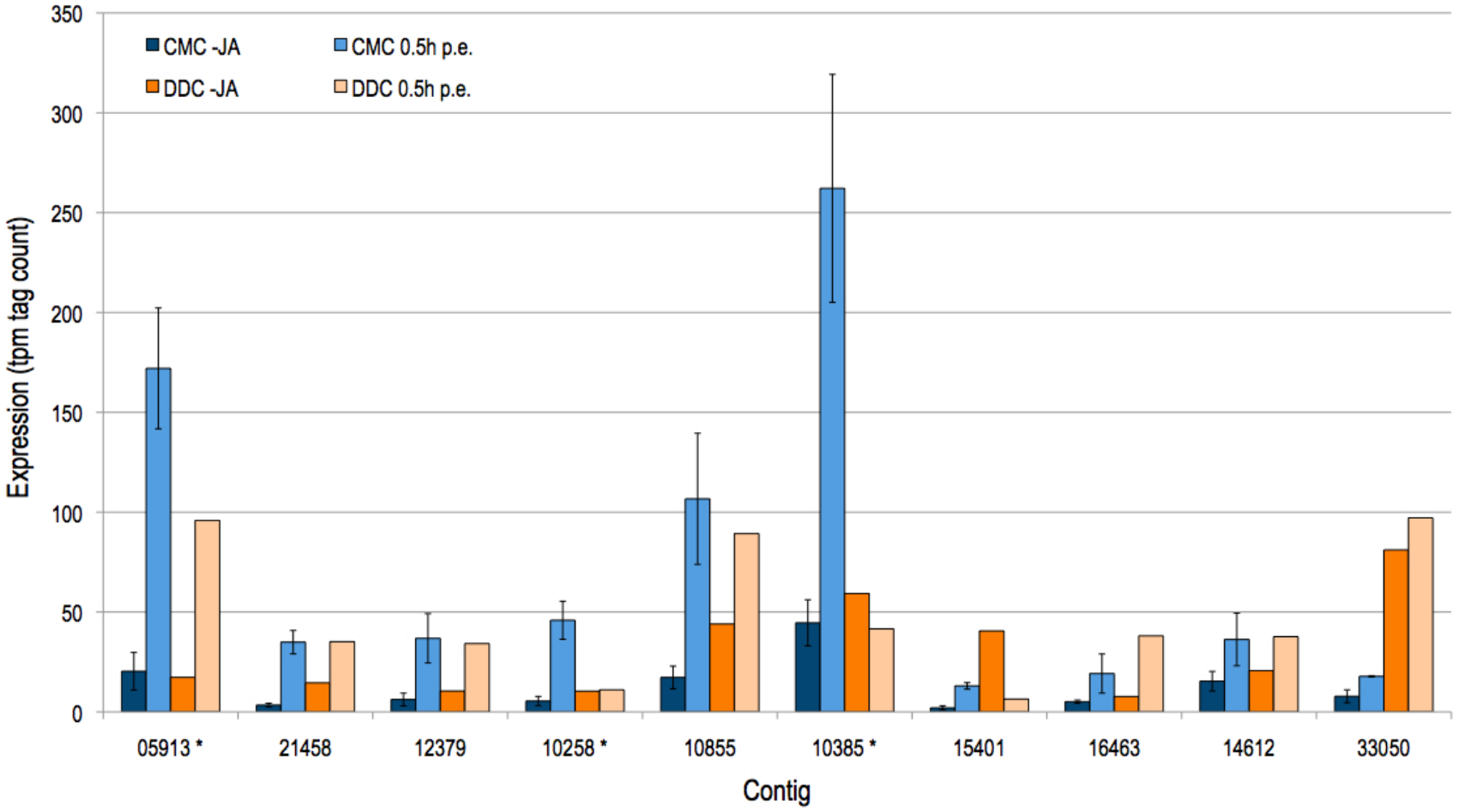

Figure 6.9: Expression levels of identified MYB TFs in CMCs and DDCs which show a significant fold change at **0.5 hours p.e.**. Asterisks indicates contigs with a differential higher expression profile in CMCs.

MYB TFs expression profiles

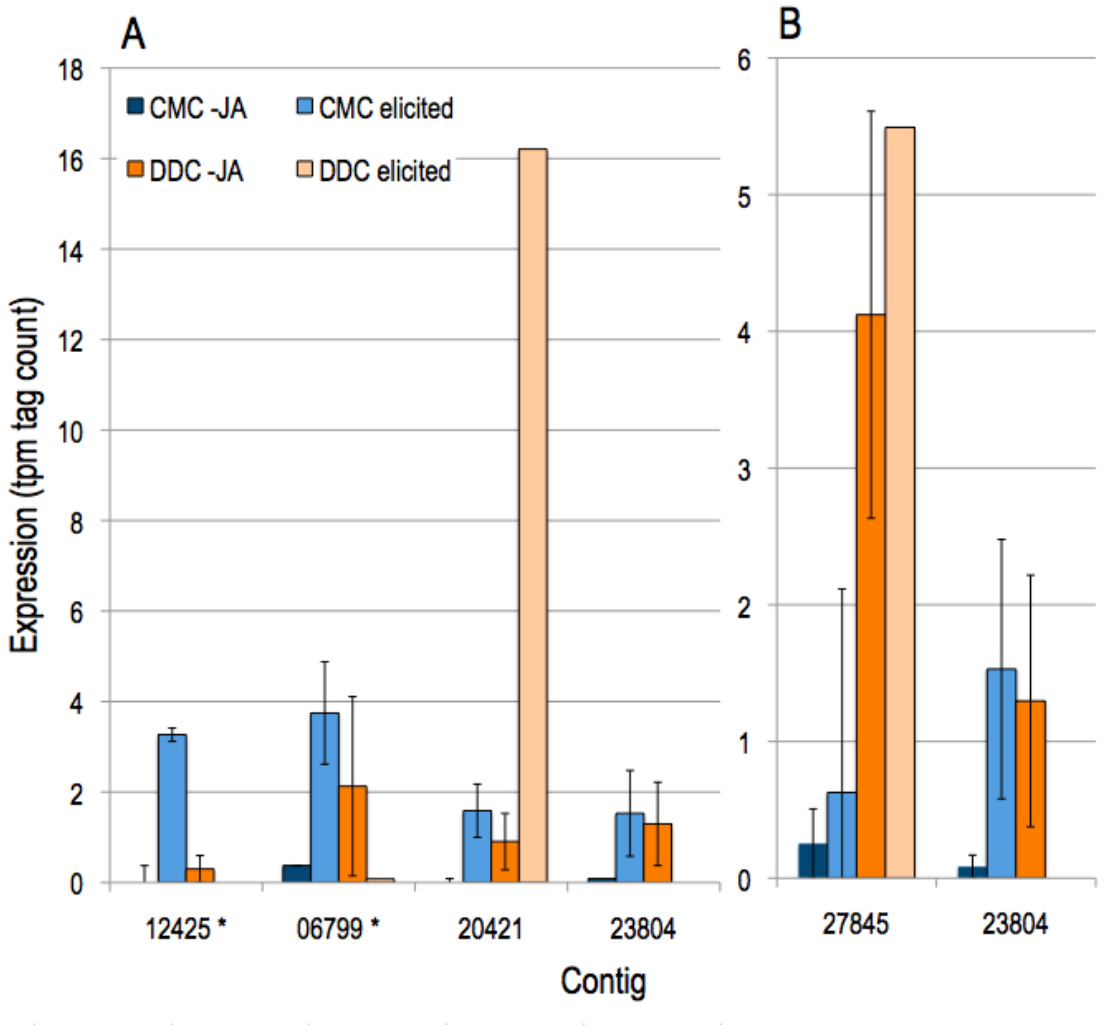

Figure 6.10: Expression levels of identified MYB TFs in CMCs and DDCs which show a significant fold change at [A]: **0.5 hours p.e.** and [B]: **2 hours p.e.** Asterisks indicates contigs with a differential higher expression profile in CMCs.

bZIP TFs expression profiles

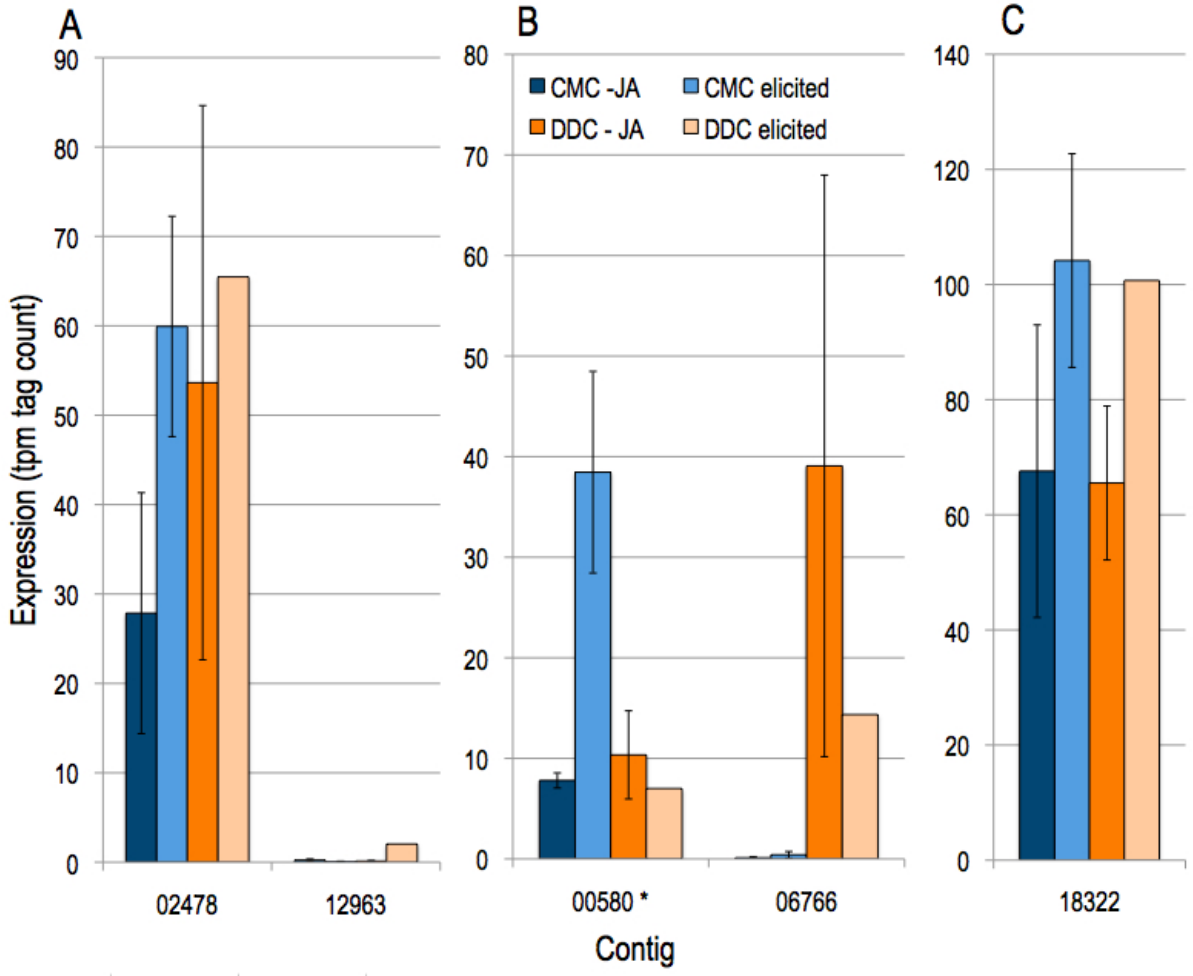

Figure 6.11: Expression levels of identified bZIP TFs in CMCs and DDCs which show a significant fold change at [A]: **0.5 hours p.e.**, [B]: **2 hours p.e.** and [C]: **12 hours p.e.** Asterisks indicates contigs with a differential higher expression profile in CMCs.

Appendix VII: *A. thaliana* homologes and function of target TFs identified in *T. cuspidata*

TF family	contig	*A. thaliana* homologue	E-value	Function	
AP2/ERF	00499	At5g44210	1×10^{-22}	Pathogenesis related TF	[Camehl and Oelmueller, 2010]
	22386	At5g13330	4×10^{-25}	Developmental and stress signalling	[Krishnaswamy et al., 2011]
	14838	n/a	n/a	n/a	n/a
	26170	At5g25810	4×10^{-25}	Multidimensional cell growth	TAIR
	07245	n/a	n/a	n/a	n/a
	03304	n/a	n/a	n/a	n/a
	17139	AT3G54320	4×10^{-34}	AP2 class, storage compound biosynthesis	TAIR
	01431	AT5G61890	2×10^{-26}	n/a	n/a
	03038	n/a	n/a	n/a	n/a
	17529	n/a	n/a	n/a	n/a
bHLH	08058	AT1G32640	3×10^{-28}	JA mediated regulation of flavonoid biosynthesis	TAIR
	15240	AAL55713	2×10^{-60}	Regulation of anthocyanin biosynthesis	[Nakatsukaa et al., 2009]
WRKY	27015	AT2G30590	5×10^{-61}	Salicylic acid inducible TF	[Tsutsui et al., 2006]
	26745	n/a	n/a	n/a	n/a
	20584	n/a	n/a	n/a	n/a
MYB	05913	n/a	n/a	n/a	n/a
	10258	n/a	n/a	n/a	n/a
	10385	AT3G23250	3×10^{-58}	Defence responsive TF	[Libault et al., 2007]
	12425	AT5G14750	1×10^{-22}	root and hypocotyl developmental process	TAIR
	06799	n/a	n/a	n/a	n/a
bZIP	00580	AT1G75390	2×10^{-24}	Stress responsive TF	[Weltmeier et al., 2009]

Appendix VII: Taxol pathway genes

Gene	Short	Accession	Reference	Contig
TAXADIENE SYNTHASE	*TASY*	DQ305407	[Wildung and Croteau, 1996]	11265
TAXADIENE 5-ALPHA HYDROXYLASE	*T5αH*	AY289209	[Jennewein et al., 2004]	14751
TAXANE 13-ALPHA HYDROXYLASE	*T5H*	AY056019	[Jennewein et al., 2001]	01720
2-DEBENZOYL-7,13-DIACETYLBACCATIN III-2-O-DEBENZOYL TRANSFERASE	*DBBT*	AF297618	[Walker and Croteau, 2000a]	04884
10-DEACETYLBACCATIN III-10-O-ACETYL TRANSFERASE	*DBAT*	AF193765	[Walker and Croteau, 2000b]	01477
BACCATIN III PHENYLPROPANOYLTRANSFERASE	*BAPT*	AY082804	[Walker et al., 2002]	17720
3'-N-DEBENZOYLTAXOL N-BENZOYLTRANSFERASE	*DBTNBT*	AF466397	[Walker et al., 2002]	03227
PHENYLALANINE AMINOMUTASE	*PAM*	AY582743	[Jennewein et al., 2004]	02897

nature biotechnology

Cultured cambial meristematic cells as a source of plant natural products

Eun-Kyong Lee[1,5], Young-Woo Jin[1,5], Joong Hyun Park[1], Young Mi Yoo[1], Sun Mi Hong[1], Rabia Amir[2], Zejun Yan[2], Eunjung Kwon[2,3], Alistair Elfick[3], Simon Tomlinson[4], Florian Halbritter[4], Thomas Waibel[2], Byung-Wook Yun[2] & Gary J Loake[2]

A plethora of important, chemically diverse natural products are derived from plants[1]. In principle, plant cell culture offers an attractive option for producing many of these compounds[2,3]. However, it is often not commercially viable because of difficulties associated with culturing dedifferentiated plant cells (DDCs) on an industrial scale[3]. To bypass the dedifferentiation step, we isolated and cultured innately undifferentiated cambial meristematic cells (CMCs). Using a combination of deep sequencing technologies, we identified marker genes and transcriptional programs consistent with a stem cell identity. This notion was further supported by the morphology of CMCs, their hypersensitivity to γ-irradiation and radiomimetic drugs and their ability to differentiate at high frequency. Suspension culture of CMCs derived from *Taxus cuspidata*, the source of the key anticancer drug, paclitaxel (Taxol)[2,3], circumvented obstacles routinely associated with the commercial growth of DDCs. These cells may provide a cost-effective and environmentally friendly platform for sustainable production of a variety of important plant natural products.

Only plant stem cells, embedded in meristems located at the tips of shoots and roots or contained inside the vascular system, can divide and give rise to cells that ultimately undergo differentiation while simultaneously giving rise to new stem cells[4]. These cells can be considered immortal due to their ability to theoretically divide an unlimited number of times. Consequently, since the beginnings of tissue culture in the 1940s, cell suspension cultures have been routinely generated through what was believed to be a dedifferentiation process[5]. Recent evidence suggests this mechanism might not entail a simple reverse reprogramming[6]. Regardless of the mechanism involved, this process results in mitotic reactivation of specialized cell types within a given organ, generating a multicellular mixture of proliferating cells[7]. Suspension cultures derived from such cellular assortments often exhibit poor growth properties with low and inconsistent yields of natural products[3], owing to deleterious genetic and epigenetic changes that occur during this process[7,8].

To circumvent this so-called dedifferentiation procedure, we developed an innately undifferentiated cell line derived from cambium cells, which function as vascular stem cells[9]. Also, paclitaxel biosynthesis in *T. cuspidata* is most conspicuous within the region containing these CMCs[10]. A recently developed twig was collected from a wild yew, *T. cuspidata* (**Fig. 1a**). We gently peeled tissue that contained cambium, phloem, cortex and epidermis from the xylem (**Fig. 1b** and **Supplementary Fig. 1a–c**) and confirmed the absence of xylem cells by staining with phloroglucinol-HCl, which detects lignin deposition (**Supplementary Fig. 2a–f**). After this tissue was cultured on solid isolation medium for 30 d (**Fig. 1c**), actively proliferating cambium cells could be gently separated from the DDCs derived from phloem, cortex and epidermis (**Fig. 1c–e** and **Supplementary Fig. 3a–e**). This mass of proliferating cells was distinct from DDCs derived from a needle or embryo (**Fig. 1f,g**), and the morphology of these CMCs differed from adjacent cells (**Fig. 1h** and **Supplementary Fig. 3b–e**). We also used this technology to produce such cells from a variety of plant species, including ginseng (*Panax ginseng*), ginkgo (*Ginkgo biloba*) and tomato (*Solanum lycopersicon*). This suggests that the procedure has broad utility (**Supplementary Fig. 4a–f**).

Microscopic analysis of a suspension culture of *T. cuspidata* cells revealed the presence of small, abundant vacuoles within the cultured cells. This characteristic feature of CMCs[11] enables them to withstand the pressure generated by the expanding secondary xylem[12]. In contrast, dedifferentiated *T. cuspidata* cells derived from needles or embryos possessed only one large vacuole, typical of such plant cells (**Fig. 1i,j**). The ability to differentiate into either a tracheary element, the main conductive cell of the xylem, or a phloem element is a defining trait of CMCs[13,14]. These cultured cells could be conditionally differentiated into a tracheary element at high frequency. In contrast, no tracheary elements were formed from *T. cuspidata* DDCs (**Fig. 1k,l**). Both animal and plant stem cells are particularly sensitive to cell death triggered by ionizing radiation, to safeguard genome integrity in populations of such cells[15]. In a similar fashion, these cultured cells are hypersensitive to γ-irradiation (**Fig. 1m**) and display increased cell death in response to the radiomimetic drug zeocin[15] (**Fig. 1n**). In aggregate, our findings, based on a variety of

[1]Unhwa Corp., Wooah-Dong, Dukjin-gu, Jeonju, South Korea. [2]Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, King's Buildings, Edinburgh, UK. [3]School of Engineering, University of Edinburgh, King's Buildings, Edinburgh, UK. [4]Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, King's Buildings, Edinburgh, UK. [5]These authors contributed equally to this work. Correspondence should be addressed to G.J.L. (gloake@ed.ac.uk).

Received 24 August; accepted 27 September; published online 24 October 2010; doi:10.1038/nbt.1693

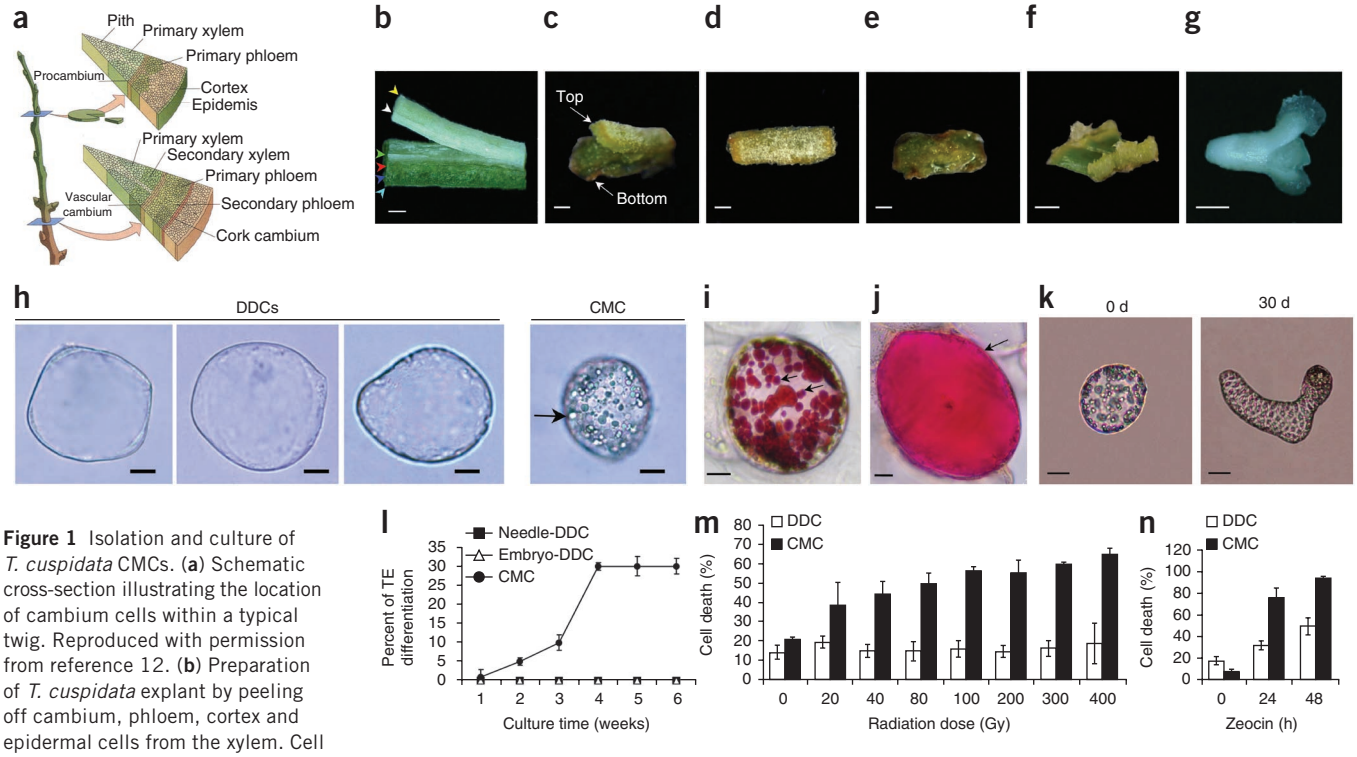

Figure 1 Isolation and culture of *T. cuspidata* CMCs. (**a**) Schematic cross-section illustrating the location of cambium cells within a typical twig. Reproduced with permission from reference 12. (**b**) Preparation of *T. cuspidata* explant by peeling off cambium, phloem, cortex and epidermal cells from the xylem. Cell types are indicated by the following colored arrows: yellow, pith; white, xylem; green, cambium; red, phloem; blue, cortex; and turquoise, epidermis. Scale bar, 0.5 mm. (**c**) Natural split of CMCs from DDCs induced from phloem, cortex and epidermal cells. The top layer is composed of CMCs whereas the bottom layer consists of DDCs. Scale bar, 1 mm. (**d**) CMCs proliferated from the cambium. Scale bar, 1 mm. (**e**) DDCs induced from the tissue containing phloem, cortex and epidermal cells. Scale bar, 1 mm. (**f**) DDCs induced from the cut edge of a needle explant. Scale bar, 0.5 mm. (**g**) DDCs induced from the cut edge of an embryo explant. Scale bar, 0.5 mm. (**h**) Micrographs of DDCs and a CMC. CMCs are significantly smaller and possess characteristic numerous, small vacuole-like structures. The black arrow indicates a vacuole-like structure. Scale bars, 20 µm. (**i**) Single CMC stained with neutral red, which marks the presence of vacuoles. Two of many stained vacuoles are denoted by black arrows. Scale bar, 10 µm. (**j**) Needle-derived DDC stained with neutral red. The single large vacuole present in this cell is marked by a black arrow. Scale bar, 10 µm. (**k**) Conditional differentiation of *T. cuspidata* CMCs to tracheary elements, at the times indicated, after addition of differentiation media. Scale bar, 25 µm. (**l**) Time-course of differentiation of different *T. cuspidata* cell lines over time into tracheary elements. (**m**) Quantification of cell death in *T. cuspidata* cells after exposure to increasing levels of ionizing radiation. (**n**) Levels of cell death in *T. cuspidata* cells after exposure to the radiomimetic drug, Zeocin (phleomycin). Experiments were repeated at least twice with similar results. Data points represent the mean of three samples ± s.d.

approaches, are consistent with the notion that these cultured cells exhibit stem cell–like properties, consistent with a CMC identity.

We used a combination of deep sequencing technologies to compare the molecular signatures of these cells and those of typical DDCs. First, we used an approach based on massively parallel pyrosequencing[16] to profile the *T. cuspidata* transcriptome. A total of 860,800 reads of average length 351 bp generated 301 MB of sequence (**Supplementary Fig. 5a** and **Supplementary Tables 1–3**). From these sequence data, we assembled 36,906 contigs *de novo* (average length, 700 bp; maximum length, 10,355 bp), with 8,865 contigs > 1 kb (**Supplementary Fig. 5b** and **Supplementary Tables 4–6**). We subjected contigs from the *T. cuspidata* transcriptome (**Supplementary Data Set 1**) to BLAST searches and 62% were assigned a putative function (**Supplementary Data Set 2**). This data set should provide an important resource because there is currently no large-scale sequence information derived from this division of the plant kingdom. The determination of the *T. cuspidata* transcriptome enabled us to use digital gene expression tag profiling[16] to compare gene expression in prospective CMCs with gene expression in DDCs (**Supplementary Data Set 3**) in the absence of exogenous chemical elicitors that can induce paclitaxel biosynthesis.

Digital gene expression tag profiling analysis established that 563 genes were differentially expressed in CMCs, with 296 upregulated and 267 downregulated (**Fig. 2a**, **Supplementary Figs. 6** and **7** and **Data Set 4**). A subset of these genes were validated by RT-PCR (**Supplementary Fig. 8**). *Phloem intercalated with xylem* (*PXY*) encodes a leucine-rich repeat (LRR) receptor-like kinase (RLK), which is conspicuously expressed in CMCs. *PXY* is a member of a small series of closely related LRR RLKs, mutations that impact CMC function[17]. *T. cuspidata* contig 01805 exhibits high similarity to *PXY* (**Supplementary Fig. 9a**) and is differentially expressed in our prospective CMC suspension cells (**Supplementary Data Set 4**). Analysis by qRT-PCR established that expression of contig 01805 is upregulated ninefold in these cells relative to DDCs (**Fig. 2b**).

Wooden Leg (*WOL*) encodes a two-component histidine kinase which is a member of a small gene family in *Arabidopsis*[18]. WOL-like proteins are unique in having two putative receiver or D-domains and mutations in *WOL* affect vascular morphogenesis[18]. *WOL* is expressed in the cambium[18] and *WOL*-like genes are expressed in the cambial zone of the silver birch (*Betula pendula*) and poplar (*Populus trichocarpa*)[19]. *T. cuspidata* contig 10710 exhibits high similarity to *WOL* and its related genes (**Supplementary Fig. 9b**). Gene expression analysis established that this gene is upregulated 12-fold in CMCs relative to DDCs (**Fig. 2b**).

To assess the molecular composition and the relative expression of genes indicative of given biological pathways in our prospective CMCs, we performed enrichment analysis of Gene Ontology (GO) terms within our data set. This approach established that both stress and biotic defense response genes were prominently over-represented

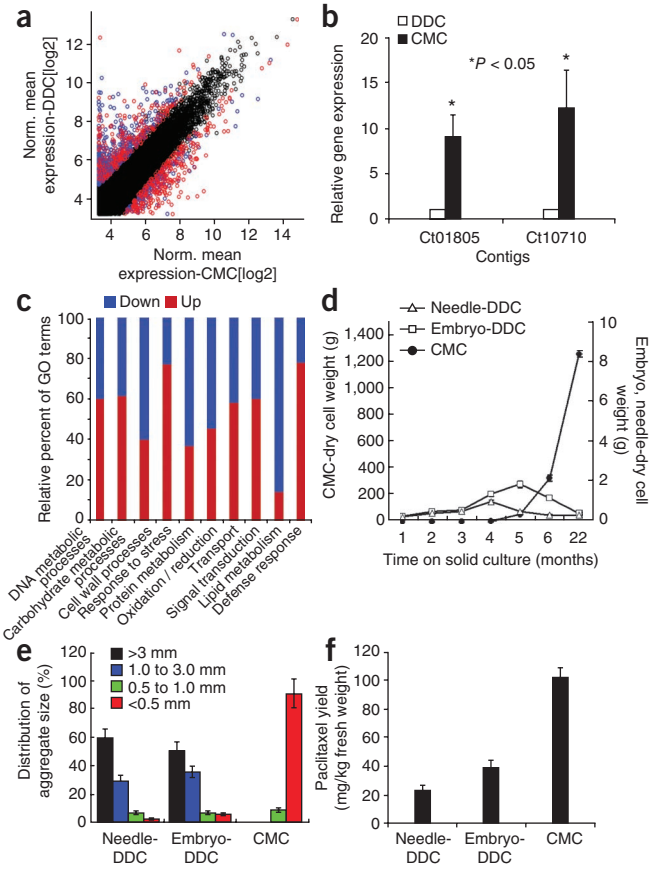

Figure 2 Characterization of CMCs from *T. cuspidata*, including transcriptome profiling using digital gene expression tags. (**a**) Scatter plot showing differentially expressed genes (DEGs) (blue and red) in CMCs and non-DEGs (black). The deployment of further filtering approaches identified more robust DEGs (red), whereas other DEGs (blue) were filtered out. FDR ≤ 0.05; $n = 1,229$. (**b**) Analysis of the expression of contig 01805 and contig 10710. (**c**) Relative percentage of GO groups within CMC DEGs. (**d**) Growth of CMCs and DDCs derived from needles or embryos on solid growth media from an initial 3 g f.c.w. 95% confidence limits are too small to be visible on this scale. (**e**) Bar graph reporting the extent of cell aggregation in DDC and CMC suspension cultures. (**f**) Paclitaxel production by 3-month-old DDCs and CMCs 10 d after elicitation, following batch culture in a flask format. Error bars represent 95% confidence limits. These experiments were repeated three times with similar results.

(**Fig. 2c**). Stem cells exhibit a low threshold for auto-execution through apoptosis but express robust defenses against environmental stresses[20]. Collectively, our Digital gene expression tag profiling data are therefore consistent with a CMC identity for these cultured cells.

We used a solid growth media format and representative cell lines to compare the growth properties of these CMCs with DDCs derived from the same wild tree. At 22 months after inoculation, the *T. cuspidata* needle- and embryo-derived DDCs produced a total dry cell weight (d.c.w.) of 0.32 g and 0.41 g, respectively, when grown on solid media with subculturing every 2 weeks (**Fig. 2d**). In contrast, the d.c.w. generated from CMCs was 1,250 g, an increase of 4.0×10^5% and 3.0×10^5%, respectively. Moreover, these cells were still growing rapidly following 22 months of culture, whereas DDCs possessed conspicuous necrotic patches (**Supplementary Fig. 10**) and displayed signs of a rapid decrease in their viable cell biomass.

Pronounced cell aggregation is a typical feature of suspension cultures comprised of DDCs. This can lead to differences in local environments between cells significantly reducing growth rate and natural product biosynthesis[3]. In representative suspension cultures of DDCs derived from either *T. cuspidata* needles or embryos, the proportion of cell aggregates with a diameter <0.5 mm was only 2% or 5%, respectively ($n = 150$ cells). Conversely, in representative CMCs, 93% of cell aggregates had a diameter <0.5 mm, with many cells present as singletons or components of aggregates comprised of only 2-3 cells ($n = 150$ cells) (**Fig. 2e** and **Supplementary Fig. 11**).

We next determined the magnitude of paclitaxel biosynthesis in this novel cell line during batch culture in a 125 ml Erlenmeyer flask. At 14 d after inoculation of flask cultures with cells, cells were transferred to production medium containing the elicitors methyl-jasmonate[2] and chitosan, together with a precursor phenylalanine, to induce paclitaxel biosynthesis. Levels of paclitaxel were measured 10 d later by high-performance liquid chromatography (HPLC). The amount of paclitaxel produced, 102 mg/kg fresh cell weight (f.c.w.), was conspicuously greater than that generated by either needle or embryo-derived DDCs at a f.c.w. value of 23 mg/kg or 39 mg/kg, respectively (**Fig. 2f**). Measurements of this natural product were confirmed by liquid chromatography mass spectrometry (LC-MS) (**Supplementary Fig. 12a–d**). Further, genes encoding key enzymes integral to the biosynthesis of paclitaxel[2,3] were induced more strongly in CMCs than in DDCs (**Supplementary Fig. 13**).

To establish whether these cells exhibit superior growth properties on a larger scale, we first investigated their performance in a 10 liter stirred tank bioreactor. In this environment, shear stress can limit growth, and the problem is often intensified by cell aggregation[21]. The CMCs in this bioreactor grew significantly faster than DDCs (**Fig. 3a**). Further, in response to shear stress, the survival rate of CMCs was strikingly higher than for DDCs, which by the end of the culture period had largely turned necrotic and had stopped growing (**Supplementary Fig. 14a–c**).

Next, we explored the performance of these cells in a 3 liter air-lift bioreactor. Large aggregates of DDCs formed in this bioreactor, leading to reduced cell mixing and circulation, which subsequently resulted in cell adherence to the bioreactor wall. Furthermore, many of these adhered cellular aggregates developed necrotic patches. After 4 months of culture, the growth of DDCs from either needle or embryo, expressed as dry cell weight (d.c.w.), were 3.33 g and 5.08 g, respectively. In contrast, the CMC line had generated a d.c.w. of 3,819.44 g, an increase of 114,000% and 75,000%, respectively (**Fig. 3b**). We also analyzed the growth of these cell lines in a 20 liter air-lift bioreactor, routinely used as a pilot for subsequent large-scale production. DDCs did not grow in this size bioreactor under the conditions tested and rapidly became necrotic. Conversely, CMCs always grew rapidly, increasing their d.c.w. from 3.65 g/l to 12.85 g/l within 14 d (**Fig. 3c**). Their relative tolerance of shear stress can likely be attributed to their small and abundant vacuoles, reduced aggregation and thin cell walls[21].

We attempted to improve the performance of needle- and embryo-derived DDCs by specifically selecting the more rapidly growing cells at each subculture on solid media for a period of 1.8 years. This process improved the growth of needle-, but not embryo-derived, DDCs in a 3 liter air-lift bioreactor. However, the performance of CMCs was still strikingly superior to that of these extensively selected cells with respect to specific growth rate (μ), doubling time (T_d) and growth index (GI) (**Supplementary Fig. 15** and **Supplementary Table 7**). A key trait for the exploitation of plant cells on an industrial scale is the stability of their growth in suspension culture[3]. We therefore monitored the growth stability of these cells compared to selected

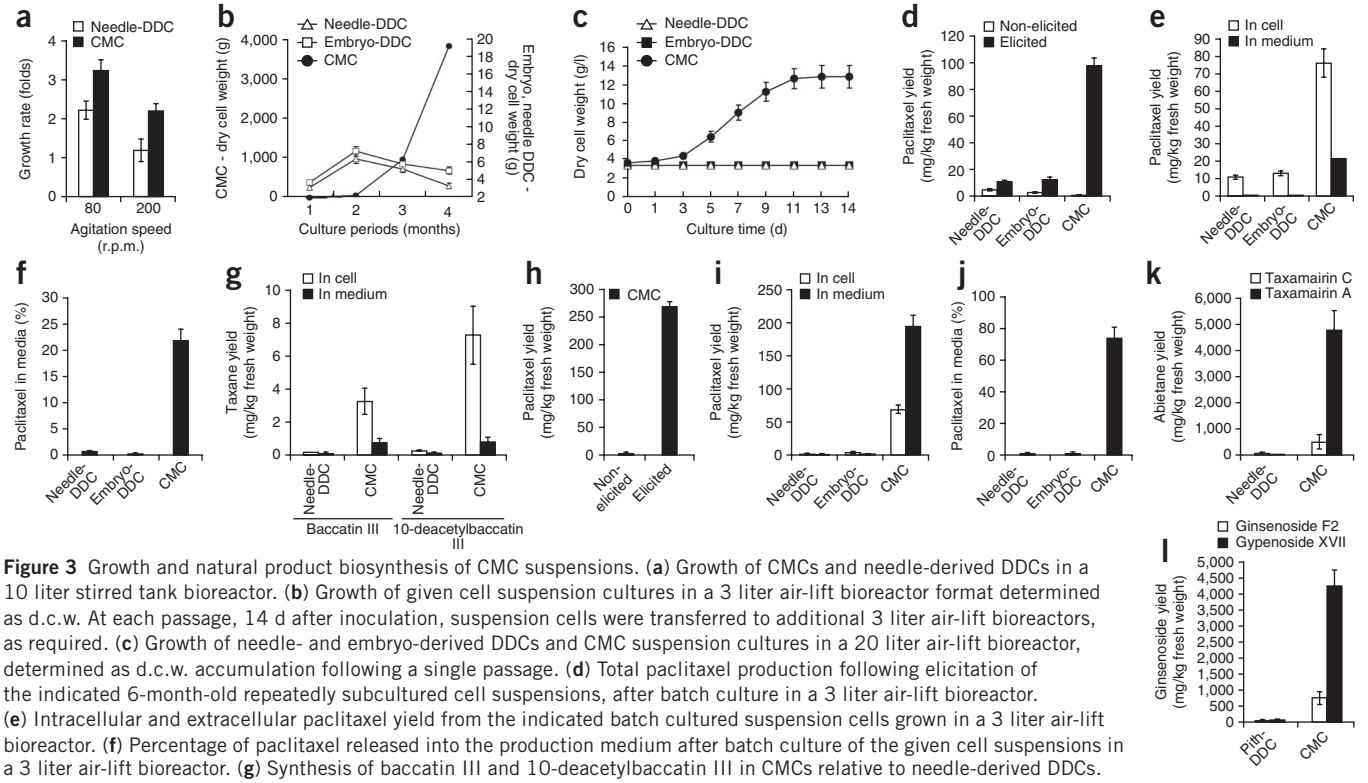

Figure 3 Growth and natural product biosynthesis of CMC suspensions. (**a**) Growth of CMCs and needle-derived DDCs in a 10 liter stirred tank bioreactor. (**b**) Growth of given cell suspension cultures in a 3 liter air-lift bioreactor format determined as d.c.w. At each passage, 14 d after inoculation, suspension cells were transferred to additional 3 liter air-lift bioreactors, as required. (**c**) Growth of needle- and embryo-derived DDCs and CMC suspension cultures in a 20 liter air-lift bioreactor, determined as d.c.w. accumulation following a single passage. (**d**) Total paclitaxel production following elicitation of the indicated 6-month-old repeatedly subcultured cell suspensions, after batch culture in a 3 liter air-lift bioreactor. (**e**) Intracellular and extracellular paclitaxel yield from the indicated batch cultured suspension cells grown in a 3 liter air-lift bioreactor. (**f**) Percentage of paclitaxel released into the production medium after batch culture of the given cell suspensions in a 3 liter air-lift bioreactor. (**g**) Synthesis of baccatin III and 10-deacetylbaccatin III in CMCs relative to needle-derived DDCs. (**h**) Magnitude of paclitaxel biosynthesis following elicitation of 28-month-old CMCs in a 20 liter air-lift bioreactor. Needle- and embryo-derived DDC suspensions did not routinely grow in this size bioreactor. (**i**) Intracellular and extracellular paclitaxel yield after 45 d of perfusion of cultured needle- and embryo-derived DDCs and CMCs in a 3 liter air-lift bioreactor. (**j**) Percentage of paclitaxel released into the production medium after perfusion culture of the given cell suspensions as indicated in **i**. (**k**) Synthesis of taxamairin A and C in CMCs and needle-derived DDCs after batch culture in a 3 liter air-lift bioreactor. (**l**) Synthesis of ginsenosides in *P. ginseng* CMC and pith-derived DDC suspension cells after batch culture in a 3 liter air-lift bioreactor. Error bars represent 95% confidence limits. These experiments were repeated twice with similar results.

DDCs derived from needles. The CMCs exhibited a relatively constant growth rate over time. In contrast, we observed striking fluctuations in growth rates during the culture of DDCs (**Supplementary Fig. 16**). Finally, we determined the growth of CMCs within a 3 ton bioreactor. These cells were again successfully cultured with high performance (**Supplementary Fig. 17a** and **b**), establishing their utility for growth on an industrial scale.

We determined the level of paclitaxel production of these different *T. cuspidata* cell suspensions in both 3 liter and 20 liter air-lift bioreactors. At 10 d after elicitation, CMCs again synthesized strikingly more paclitaxel than either of the DDC lines in a 3 liter air-lift bioreactor. Further, elicitation induced a 220% (11 mg/kg) and 433% (13 mg/kg) increase in paclitaxel production within needle- and embryo-derived DDCs respectively, whereas the induction was 14,000% (98 mg/kg) with CMCs (**Fig. 3d**). CMCs secreted 2.7×10^4% and 7.2×10^4% more paclitaxel into the culture medium than the low levels secreted by either needle- or embryo-derived DDCs, respectively (**Fig. 3e**,**f**). The amount of paclitaxel secreted to the medium during culture varies significantly both between *Taxus* species and in response to different culture conditions[22]. Our DDCs secreted less paclitaxel than might be expected. Nevertheless, *T. cuspidata* CMCs secreted a strikingly greater amount of paclitaxel into the medium under these culture conditions than the associated DDCs. Moreover, these cells also synthesized strikingly more of the related taxanes baccatin III and 10-deacetylbaccatin III[2,3] (**Fig. 3g**). No paclitaxel production was detected by either DDC line in a 20 liter air-lift bioreactor. In contrast, CMCs synthesized 268 mg/kg and were again highly responsive to elicitation (**Fig. 3h**). Previously reported values for *T. cuspidata* paclitaxel production range from 20–84 mg/kg f.c.w.[23,24]. However, these data, including the maximum value, were obtained from flask cultures, whereas our data suggest that DDCs have improved function relative to their performance on a larger scale. Our findings imply that CMCs synthesize strikingly more paclitaxel and are significantly more responsive to elicitation when batch cultured in either 3 liter or 20 liter air-lift bioreactors compared to typical *T. cuspidata* suspension cells.

Perfusion culture promotes the secretion of secondary metabolites into the culture medium, aiding both purification and natural product biosynthesis[22]. We therefore compared the magnitude of paclitaxel secretion after perfusion culture. Following 45 d of perfusion culture, needle- and embryo-derived DDCs were largely necrotic; however, CMCs produced a combined total of 264 mg of paclitaxel per kg of cells and 74% of this was secreted directly into the medium (**Fig. 3i**,**j**). Perfusion culture of these cells therefore both promotes paclitaxel biosynthesis and increases the proportion of this secondary product that is secreted into the medium, facilitating its cost-effective purification. The future deployment of metabolic engineering approaches and higher yielding *Taxus* species may further enhance paclitaxel biosynthesis in these cells[2,3].

We also monitored these *T. cuspidata* suspension cultures for the production of the abietane tricyclic diterpenoid derivatives, taxamairin A and taxamairin C, which have also been shown to possess antitumor activities[25]. Elicitation of these cells within a 3 liter air-lift bioreactor induced increases in both taxamairin C and especially taxamairin A to 520.8 and 4,982.5 mg/kg f.c.w., respectively, in CMCs. These values were far greater than those determined in

DDCs (**Fig. 3k**). Suspension cultures of *T. cuspidata* have previously been reported to produce 0.92 and 26.08 mg/kg f.c.w. of taxamairin C and taxamairin A, respectively[26]. Our data imply that CMCs might provide a considerably better source of these abietanes than DDCs. To establish whether CMCs derived from other plant species also exhibit superior properties with respect to the biosynthesis of commercially relevant natural products, we determined the synthesis of ginsenosides, a class of triterpenoid saponins derived exclusively from the plant genus *Panax*. Ginsenosides have been reported to show multiple bioactivities including neuroprotection, antioxidative effects and the modulation of angiogenesis[27]. Following elicitation of tap root–derived *P. ginseng* suspension cells, cultured using a 3 liter air-lift bioreactor, ginsenoside F2 and gypenoside XVII accumulated to strikingly greater levels in *P. ginseng* CMCs relative to DDCs. Ginsenoside F2 and gypenoside XVII accrued to 791 and 4,425 mg/kg f.c.w., respectively (**Fig. 3l**). Previously, ginsenoside F2 has been reported to reach 33.3 mg/kg f.c.w.[28] and gypenoside XVII 183.3 mg/kg f.c.w.[29] in ginseng roots. Thus, CMCs synthesize 23.8- and 24.1-fold more ginsenoside F2 and gypenoside XVII, respectively, than previously described sources. Therefore, CMCs may also be used for the production of ginsenosides.

Numerous medicines, perfumes, pigments, antimicrobials and insecticides are derived from plant natural products[1–3,30]. Cultured cambial meristematic cells may provide a cost-effective, environmentally friendly and sustainable source of paclitaxel and potentially other important natural products. Unlike plant cultivation, this approach is not subject to the unpredictability caused by variation in climatic conditions or political instability in certain parts of the world. Furthermore, CMCs from reference species may also provide an important biological tool to explore plant stem cell function.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturebiotechnology/.

Accession codes. Sequence Read Archive: ERP000352.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

T.W. was awarded a BBSRC CASE PhD studentship. This project was funded in part by a grant from the Korea Institute for Advancement of Technology (KIAT) (R & D project number: 10030175), the Ministry of Knowledge Economy (MKE), Republic of Korea to E.-K.L., J.H.P., S.M.H. and G.J.L. R.A. was supported by a scholarship from HEC Pakistan. E.K. was supported by a studentship from the Engineering and Physical Sciences Research Council. We acknowledge the expert technical assistance of A. Montazam and D. Cleven for Roche 454 sequencing and M. Thomson for Illumina Solexa sequencing. Further, S. Bridgett and U. Trivedi provided invaluable input for bioinformatic analysis of the deep sequencing data. All sequencing was undertaken at the GenePool facility, University of Edinburgh.

AUTHOR CONTRIBUTIONS

E.-K.L., Y.-W.J., J.H.P., T.W. and B.-W.Y. performed experiments. R.A., E.K., S.T. and F.H. contributed to bioinformatic and statistical analysis. Z.Y., Y.M.Y. and S.M.H. carried out experiments. A.E. co-supervised E.K. E.-K.L., Y.-W.J. and G.J.L. formulated experiments. G.J.L. and E.-K.L. wrote the paper. All authors discussed results and commented on the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturebiotechnology/.

Published online at http://www.nature.com/naturebiotechnology/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Schmidt, B.M., Ribnicky, D.M., Lipsky, P.E. & Raskin, I. Revisiting the ancient concept of botanical therapeutics. *Nat. Chem. Biol.* **3**, 360–366 (2007).
2. Croteau, R., Ketchum, R.E.B., Long, R.M., Kaspera, R. & Wildung, M.R. Taxol biosynthesis and molecular genetics. *Phytochem. Rev.* **5**, 75–97 (2006).
3. Roberts, S.C. Production and engineering of terpenoids in plant cell culture. *Nat. Chem. Biol.* **3**, 387–395 (2007).
4. Laux, T. The stem cell concept in plants: a matter of debate. *Cell* **113**, 281–283 (2003).
5. Thorpe, T.A. History of plant tissue culture. *Mol. Biotechnol.* **37**, 169–180 (2007).
6. Sugimoto, K., Jiao, Y. & Meyerowitz, E.M. *Arabidopsis* regeneration from multiple tissues occurs via a root development pathway. *Dev. Cell* **18**, 463–471 (2010).
7. Grafi, G. *et al.* Histone methylation controls telomerase-independent telomere lengthening in cells undergoing dedifferentiation. *Dev. Biol.* **306**, 838–846 (2007).
8. Baebler, S. *et al.* Establishment of cell suspension cultures of yew (*Taxus x Media* Rehd) and assessment of their genomic stability. *In Vitro Cell. Dev. Biol. Plant* **41**, 338–343 (2005).
9. Ye, Z.-H. Vascular tissue differentiation and pattern formation in plants. *Annu. Rev. Plant Biol.* **53**, 183–202 (2002).
10. Strobel, G.A. *et al.* Taxol formation in yew-*Taxus. Plant Sci.* **92**, 1–12 (1993).
11. Frankenstein, C., Eckstein, D. & Schmitt, U. The onset of cambium activity - A matter of agreement? *Dendrochronologia* **23**, 57–62 (2005).
12. Moore, R., Clark, W.D., Stern, K.R. & Vodopich, D. (eds.) *Botany* (Wm.C. Brown, Dubuque, Iowa, USA; 1995).
13. Turner, S., Gallois, P. & Brown, D. Tracheary element differentiation. *Annu. Rev. Plant Biol.* **58**, 407–433 (2007).
14. Ito, Y. *et al.* Dodeca-CLE peptides as suppressors of plant stem cell differentiation. *Science* **313**, 842–845 (2006).
15. Fulcher, N. & Sablowski, R. Hypersensitivity to DNA damage in plant stem cell niches. *Proc. Natl. Acad. Sci. USA* **106**, 20984–20988 (2009).
16. Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nat. Biotechnol.* **26**, 1135–1145 (2008).
17. Fisher, K. & Turner, S. PXY, a receptor-like kinase essential for maintaining polarity during plant vascular-tissue development. *Curr. Biol.* **17**, 1061–1066 (2007).
18. Mähönen, A.P. *et al.* A novel two-component hybrid molecule regulates vascular morphogenesis of the *Arabidopsis* root. *Genes Dev.* **14**, 2938–2943 (2000).
19. Nieminen, K. *et al.* Cytokinin signaling regulates cambial development in popular. *Proc. Natl. Acad. Sci. USA* **105**, 20032–20037 (2008).
20. Rando, T.A. The immortal strand hypothesis: segregation and reconstruction. *Cell* **129**, 1239–1243 (2007).
21. Joshi, J.B., Elias, C.B. & Patole, M.S. Role of hydrodynamic shear in the cultivation of animal, plant and microbial cells. *Chem. Eng. J.* **62**, 121–141 (1996).
22. Wang, C., Wu, J. & Mei, X. Enhanced taxol production and release in *Taxus chinesis* cell suspension cultures with selected organic solvents and sucrose feeding. *Biotechnol. Prog.* **17**, 89–94 (2001).
23. Mirjalili, N. & Linden, J.C. Methyl jasmonate induced production of Taxol in suspension cultures of *Taxus cuspidata*: Ethylene interaction and induction models. *Biotechnol. Prog.* **12**, 110–118 (1996).
24. Wu, Z.L., Yuan, Y.-J., Ma, Z.-H. & Hu, Z.D. Kinetics of two-liquid-phase *Taxus cuspidata* cell culture for production of Taxol. *Biochem. Eng. J.* **5**, 137–142 (2000).
25. Yang, S.-J., Fang, J.-M. & Cheng, Y.-S. Lignans, flavonoids and phenolic derivatives from *Taxus mairei. J. Chinese Chem. Soc.* **46**, 811–818 (1999).
26. Bai, J. *et al.* Production of biologically active taxoids by a callus culture of *Taxus cuspidata. J. Nat. Prod.* **67**, 58–63 (2004).
27. Leung, K.W. & Wong, A.S.-T. Pharmacology of ginsenosides: a literature review. *Chin. Med.* **5**, 20 (2010).
28. Dan, M. *et al.* Metabolite profiling of *Panax notoginseng* using UPLC-ESI-MS. *Phytochemistry* **69**, 2237–2244 (2008).
29. Reynolds, L.B. Effects of harvest date on some chemical and physical characteristics of American ginseng (*Panax quinquefolius L.*). *J. Herbs Spices Med. Plants* **6**, 63–69 (1998).
30. Kutchan, T. & Dixon, R.A. Physiology and metabolism: Secondary metabolism: nature's chemical reservoir under deconvolution. *Curr. Opin. Plant Biol.* **8**, 227–229 (2005).

ONLINE METHODS

Collection and sterilization of *T. cuspidata* samples. Twig, needle and seed samples were collected from a wild-grown *T. cuspidata* tree. Twig and needle samples were immediately deposited in 0.56 mM ascorbic acid solution. They were stored at 4 °C for 1 month. Then, they were washed in running tap water for 30 min and surface disinfected with 70% ethanol for 1 min, followed by 1% sodium hypochlorite for 20 min for twigs and 15 min for needles and 0.07% NaOCl for 20 min, and rinsed 5 times with sterilized distilled water (dH_2O). Lastly, they were rinsed once with dH_2O containing 150 mg/l citric acid. Seeds were put into 0.01% NaOCl for 24 h with agitation. They were washed in running tap water for 4 h, surface disinfected with 70% ethanol for 1 min and then placed in a 1% NaOCl for 15 min. Then, they were rinsed 5 times with dH_2O.

Isolation of CMCs. For CMCs, cambium, phloem, cortex and epidermal tissue were peeled off from the xylem and the epidermal tissue side was laid on B5 medium[31] excluding $(NH_4)_2SO_4$ with 1 mg/l picloram, 30 g/l sucrose and 4 g/l gelrite. After 4 to 7 d, cell division was evident only in cambium and after 15 d, DDCs started to form from the layer that consisted of phloem, cortex and epidermis by dedifferentiation. At 30 d post-culture, there was a visible split between cambium cells and DDCs of the phloem, cortex and epidermis. This separation was obvious because cambium cells uniformly divided resulting in the formation of a flat plate of cells. In contrast, DDCs derived from phloem, cortex and epidermis proliferated in an irregular form, presumably due to the discrepancy between cell division rates. Following the natural separation of cambium from the other cell types, this cell layer was transferred onto different Petri dishes containing B5 medium excluding $(NH_4)_2SO_4$ with 1 mg/l picloram, 10 g/l sucrose and 4 g/l gelrite. Initial cell inoculum size was 3.0 g (f.c.w.) and subsequently, CMCs were subcultured onto the fresh medium every 2 weeks. Establishment of *P. ginseng* CMCs was as described above except that the isolation medium contained McCown woody plant medium with 2 mg/l IAA, 30 g/l sucrose, 100 mg/l ascorbic acid, 150 mg/l citric acid and 3 g/l gelrite. For lignin visualization, tissues were stained with phloroglucinol-HCl (0.5% (wt/vol) phloroglucinol in 6 N HCl) for 5 min and then observed under a light microscope.

Establishment of cell suspension cultures and natural product production. Initial suspension cultures were established by inoculating a sample of 2.5 g (f.c.w.) cultured cells derived from either cambium, needles or embryos into 125 ml Erlenmeyer flasks containing 25 ml of B5 medium containing 1 mg/l picloram, and 20 g/l sucrose, excluding $(NH_4)_2SO_4$. The flasks were agitated at 100 r.p.m. and 21 °C in the dark. Subculturing was undertaken at 2-week intervals.

For culturing the cells in 3 liter and 20 liter air-lift bioreactors, the same medium that was used in the initial suspension culture was applied. Diameter, height and pore size of micro-sparger used in the bioreactor was 2 cm, 0.4 cm, 0.2 μm, respectively. Aeration rate was 0.1–0.2 vol/vol/min (v.v.m.) in 3 liter air-lift bioreactor, and 0.08–0.18 v.v.m. in 20 liter air-lift bioreactor. 3.25 g/l (d.c.w.) of CMCs, 3.3 g/l (d.c.w.) of needle-derived DDCs and 3.1 g/l (d.c.w.) of embryo-derived DDCs were inoculated in 3 liter air-lift bioreactor. 3.65 g/l (d.c.w.) of CMCs, 3.64 g/l (d.c.w.) of needle-derived DDCs and 3.41 g/l (d.c.w.) of embryo-derived DDCs were inoculated in 20 liter air-lift bioreactor. Working volume was 80% of total capacity, which is 2.4 liters in 3 liter air-lift bioreactor and 16 liters in 20 liter air-lift bioreactor. Subculturing of CMCs and DDCs was undertaken every 2 weeks in 3 liter and 20 liter air-lift bioreactor with same initial inoculum size and conditioned medium was recycled with the ratio of 25% of working volume. Growth rate was measured in d.c.w. (g/l) after vacuum filtration and drying of the cells in an oven at 70 °C for 24 h. We call these CMCs Ddobyul, meaning 'another star' in Korean.

To test the capacity for production of paclitaxel in the flask and 3 liter and 20 liter air-lift bioreactors, cells at 14 d of culturing were transferred to B5 medium excluding KNO_3, and containing 60 g/l fructose and 2 mg/l 1-naphtalene acetic acid (NAA), and elicitors such as 50 mg/l, chitosan and 100 μM methyl-jasmonate, in addition to 0.1 mM of the precursor, phenylalanine.

Me-JA was dissolved in 90% ethanol, chitosan in glacial acetic acid and phenylalanine in distilled water before dilution to the required concentrations. After 10 d, paclitaxel content was analyzed. Taxane and abietane production was elicited in a similar fashion. Stress-triggered ginsenoside accrual was mediated by reducing air supply from a constant 0.1 v.v.m. into a 3 liter air-lift bioreactor, for 13 d of culture, to 0.1 v.v.m. for a 30 min period twice per day for 3 d.

Establishment of needle and embryo DDC cultures. DDCs were induced from embryos and needles largely as previously described[32,33]. For induction of needle-derived DDCs, both ends of the needle were cut in 0.3~0.5 cm (length and width) and were laid on B5 medium containing 1 mg/l picloram, 30 g/l sucrose and 4 g/l gelrite, excluding $(NH_4)_2SO_4$. After 30 d of culturing, DDCs were induced from the cut edges (**Fig. 1f**). As culture period continued, DDCs formed over the whole explants. Induced DDCs were transferred to B5 medium containing 1 mg/l picloram, 10 g/l sucrose and 4 g/l gelrite, excluding $(NH_4)_2SO_4$ for growth. Initial inoculum size was 3.0 g (f.c.w.) and DDCs were subcultured to fresh medium every 2 weeks.

For induction of embryo-derived DDCs, both ends of the zygotic embryo were cut and laid on B5 medium containing 1 mg/l 2,4-D, 30 g/l sucrose and 4 g/l gelrite, excluding $(NH_4)_2SO_4$. After 23 d of culturing, DDCs were induced from the cut edges (**Fig. 1g**). As culture period continued, DDCs formed over the whole explant. Induced DDCs were transferred to B5 medium containing 1 mg/l picloram, 10 g/l sucrose and 4 g/l gelrite, excluding $(NH_4)_2SO_4$ for growth. Initial inoculum size was 3.0 g (f.c.w.) and DDCs were subcultured to fresh medium every 2 weeks.

CMC differentiation. The media used to induce CMC differentiation into tracheary elements (TEs) was B5 medium, excluding $(NH_4)_2SO_4$ with 10 mg/l NAA, 2 mg/l kinetin, 6 mg/l GA_3 and 60 g/l sucrose. TEs were identified by virtue of their bright fluorescence, due to the presence of lignified secondary cell walls. The extent of TE differentiation was determined as the percentage of TEs per total number of cells. This analysis was undertaken in triplicate and in each case 200 cells were counted.

Response of *T. cuspidata* cell suspensions to γ-irradiation and radiomimetic drugs. CMCs and needle-derived DDCs were obtained from suspension cultures obtained from 20 liter air-lift bioreactors. For gamma-irradiation (Co^{60}), cells were irradiated at a dose rate of 0.92 Gy/min for 0~400 Gy, which has been modified from a method described previously[34]. Then, cells were suspension cultured for 24 h in 100 ml flasks at 21 °C, 100 r.p.m. in the dark (volume of cells to liters of medium was 1:10). Suspension cells were treated with Zeocin (200 μg/ml, Invitrogen) at 7 d after subculture, essentially as described previously[34]. The treated suspension cell culture was incubated in the dark for 24 or 48 h. For cell death determination, cells were treated with 2% Evan's blue for 3 min and washed with sterile water several times, then transferred to a microscope slide covered with a thin cover slip. For each sample, cell death was determined 5 times independently and the average cell death rate was measured by excluding the maximum and minimum number of cell counts. All experiments were undertaken in triplicate.

Determination of *T. cuspidata* transcriptome. RNA was isolated using a Qiagen plant RNA kit following the manufacturer's instructions. cDNA was synthesized by employing a SMART procedure to enrich for full-length sequences[35]. The resulting cDNA was normalized using kamchatka crab duplex-specific nuclease[36], to aid the discovery of rare transcripts. cDNA was sheared using Covaris instruments settings: target size 500 bp, duty cycle 5%, intensity 3, cycles/burst 200 and time 90 s. Library preparation was undertaken using a Roche GS FLX library kit. The concentration and quality of the synthesized library was analyzed using a Agilment bio-analyzer. Titration emulsion PCR using a GS FLX emPCR kit was undertaken to determine the optimum number of beads to load for large-scale sequencing. A Beckman/Coulter Multisizer 3 bead counter was employed to determine the concentration of beads. Two million beads were loaded onto a GS FLX pico titer plate using a Roche 70 × 75 kit.

The sequencing reagents used and washes undertaken followed protocols from the manufacturer. The *T. cuspidata* transcriptome was determined in the GenePool genomics facility at the University of Edinburgh using a Roche 454 GS FLX instrument in titanium mode, which uses massively parallel pyrosequencing technology[37,38]. A total of 860,800 reads were achieved of 351 bp average length, which generated 301 MB of sequence. These data were

doi:10.1038/nbt.1693

assembled into isotigs by employing Newbler 2.3. BLAST was used to search for similar sequences within available sequence databases. Annot8r was employed to predict GO terms for each isotig[39].

Digital gene expression tag profiling. The analysis of global gene expression in *T. cuspidata* cell suspension cultures was carried out by digital gene expression tag profiling, using an improved method based on previously described technology[40]. Potentially contaminating DNA was removed from RNA samples using Ambion turbo DNase treatment. NlaIII library preparation was accomplished by following the manufacturer's standard protocol. Fifteen PCR cycles were used for amplification. We used 1–10 µg of a given library for sequencing from each sample. Sequencing was carried out in the GenePool genomics facility at the University of Edinburgh using a Genome Analyser (GA) II_x Illumina Solexa sequencing machine. Three replicates each of both cell lines (CMC/DDC) were sequenced following the manufacturer's protocol. Subsequently, reads were aligned to the *T. cuspidata* reference genome using MAQ v. 6.0.8. and uniquely aligned reads to the previously assembled *T. cuspidata* transcriptome were counted.

Statistical analysis. Statistical analysis was performed in R using the edgeR Bioconductor library[40,41]. We sought to reduce problems created by varying library sizes and noise for genes not highly expressed, by stabilizing read counts through adding a small constant. Therefore we first rescaled the read counts in each library by dividing by the sum of all read counts in the upper quartile of expression values[42] and afterwards added a constant factor C (C = 10) to each count. This transformation alters the signal in such a way that differences between groups for contigs with low expression are less likely to be considered differentially expressed, while leaving high transcript counts largely untouched. Briefly, edgeR uses an overdispersed Poisson distribution to model read count data, where the degree of overdispersion is moderated using an empirical Bayes procedure. Differential expression is assessed using a modified version of Fisher's exact test. We ran edgeR according to the steps outlined in the library's tutorial (using parameter settings prior. $n = 10$, grid.length = 500). P-values were adjusted for the false discovery rate and we deemed a threshold of false discovery rate (FDR) ≤ 0.05 to be appropriate to detect differentially expressed contigs ($n = 1{,}229$).

In the latter analysis, we decided to first focus on only those differentially expressed contigs, that showed a considerably large change (that is, the minimum difference between any replicates in both groups, DDC and CMC, was at least ten transcripts per million) and for which the direction of change was consistent between all replicates (that is, all replicates are either higher or all replicates are lower in one group than in the other). We considered these filtered contigs ($n = 563$) the most interesting candidates for immediate study and held out the rest for further follow-up studies.

Gene expression analysis. The determination of gene expression levels were carried out by either RT-PCR or qRT-PCR as previously described[43]. The primer sequences employed are listed below:

Primer sequences for qRT-PCR.
Ct01805-F: CTTGGCAAGGATCCAGTTTAG
Ct01805-R: AGACCAAGCCCAGGGTCTTC
Ct10710-F: TTCTTCGGCTGTCAGTGATG
Ct10710-R: CCGATAGAAGCTTGCAGGAA
Primer sequences for RT-PCR.
Ct27072-F: CACTTGGAGTTCGTCGTTGA
Ct27072-R: CACTGTGCACACTCACCAAA
Ct36802-F: GAGCCGTTGCATGGTACACT
Ct36802-R: TAACCGTGGTGCTCAAATCA
Ct18649-F: CCTGACAACAGCGTCTCTGA
Ct18649-R: AAACCACCAGTACCCACAGC
Ct33753-F: GTTAGACCCTTCACCGTCCA
Ct33753-R: CTGCAAAGATGAGAGTGGAATG
Ct30863-F: GCAACGTCTGAAACGCAGTA
Ct30863-R: AGAGTTGCGAACAGCAAAGG
Ct34959-F: ACTCGATAGAGCCGACAAGG
Ct34959-R: CAGCTGATCGTCCAGCTATG
Ct01720-F: CTCCTCTCCAACGAGGAAAA
Ct01720-R: GTTTTCCCCAGAAGGGAATC
Ct09814-F: TTTGAGGCATGTGGGTTTTA
Ct09814-R: TGTCAATCTGTTGCATTGGA
Ct07968-F: CGACAACATTCTTGCATTGA
Ct07968-R: AACCGTTGCAGGGAACTTAC
Ct03409-F: ATGTTCCAAAAATGGGAGGA
Ct03409-R: GCTTGGAAAGACCTGAAGGA
Ct04884-F: AGTGAATGTAAGCCCCATGA
Ct04884-R: TTTGGCATCTTCTTGGATGA
Ct07286-F: GTCCATCCATTGTCCATAGAAA
Ct07286-R: TGGCAACATTGGTAAAGATATTCA

Perfusion culture. Perfusion culture was initiated in a similar fashion to that described for the bioreactors. On day 14, cultures were elicited with 50 mg/l chitosan, 0.1 mM phenylalanine and 100 µM methyl jasmonate. After elicitation, the spent medium was removed aseptically and replaced with an equal volume of fresh B5 medium excluding KNO_3 with 60 g/l fructose and 2 mg/l 1-naphtalene acetic acid (NAA) and elicitors of 50 mg/l chitosan, 0.1 mM phenylalanine and 100 µM methyl-jasmonate every 5 d. After 45 d of extended culture, intracellular and extracellular paclitaxel levels were analyzed.

Analysis of taxanes (paclitaxel, baccatin III, 10-deacetylbaccatin III) content. After their separation from the production medium, 0.2 g of cells were weighed, soaked in 4 ml of methanol (Sigma)/dichloromethane (Sigma) (1:1 vol/vol) and sonicated (Branson) for 1 h. The methanol/dichloromethane extract (4 ml) was filtered and concentrated *in vacuo* and subsequently redissolved in 4 ml of dichloromethane and partitioned with 2 ml of water. The latter step was repeated three times and only the dichloromethane fraction was collected. This fraction was concentrated, then redissolved in 1 ml of methanol and centrifuged at 8,000*g* for 3 min before HPLC analysis. For determining the extracellular paclitaxel concentration, production medium (5 ml) was extracted 3 times with the same volume of dichloromethane. The combined dichloromethane fraction was subsequently concentrated and then redissolved in 0.5 ml methanol. HPLC (nanospace SI-2, Shiseido) with a C18 column (Capcell pak C18 MGII column, 5 µm, 3.0 mm × 250 mm, Shiseido) was used for the analysis. Column temperature was 40 °C and the mobile phase was a mixture of water and acetonitrile (Burdick & Jackson) (1:1 isocratic) at a flow rate of 0.5 ml/min. A UV-VIS detector monitored at 227 nm and the sample injection volume was 10 µl. Authentic paclitaxel, baccatin III, 10-deacetylbaccatin III standard was purchased from Sigma.

Analysis of abietanes (taxamairin A and taxamairin C) content. After their separation from the production medium, 20 mg of lyophilized cells were weighed, soaked in 4 ml of methanol (Sigma)/dichloromethane (Sigma) (1:1 vol/vol) and sonicated (Branson) for 1 h. The methanol/dichloromethane extract (4 ml) was filtered and concentrated *in vacuo* and subsequently redissolved in 4 ml of dichloromethane and partitioned with 2 ml of water. The latter step was repeated three times and only the dichloromethane fraction was collected. This fraction was concentrated, then redissolved in 1 ml of methanol and centrifuged at 8,000*g* for 3 min. Then it was filtered through 0.2 µm filter for UPLC analysis. UPLC (Waters) with a C18 column (BEH C18 1.7 µm, 2.1 × 100 mm Waters) was used for the analysis. Column temperature was 40 °C and the mobile phase was a mixture of water and acetonitrile (Burdick & Jackson) flow rate of 0.4 ml/min. Water (solvent A) and acetonitrile (solvent B) as mobile phase with a linear gradient was used: (1 min: 0% B, 13 min: 100% B, 15 min: 100% B, 16.2 min: 0% B, 17 min: 0% B). A UV-VIS detector monitored at 210 nm and the sample injection volume was 2 µl. Authentic taxamairin A and taxamairin C standard were isolated at Unhwa.

Analysis of ginsenosides (ginsenoside F2, gypenoside XVII) content. Compounds of *Panax ginseng* CMCs were analyzed through HPLC-ELSD (Younglin) and two major peaks were isolated. The two compounds isolated were identified as ginsenoside F2 and gypenoside XVII through LC-MS (Agilent), ^{1}H NMR, ^{13}C NMR and 2D NMR (Varian). For quantification of ginsenoside F2 and gypenoside XVII in *Panax ginseng* CMCs, cultured cells

were separated from the medium and were lyophilized. 100 mg of lyophilized cells were put into 2 ml of methanol (Sigma), vortexed for 5 min, and were extracted for 1 h. Cells were centrifuged at 8,000*g* for 3 min. After concentration of the supernatant, it was dissolved in 200 µl of methanol and filtered through 0.2 µm filter for UPLC analysis. UPLC with a C18 column (BEH C18 1.7 µm, 2.1 × 100 mm Waters) was used for the analysis. Column temperature was 40 °C and the mobile phase was a mixture of water and acetonitrile (Burdick & Jackson), flow rate of 0.4 ml/min. Water (solvent A) and acetonitrile (solvent B) as mobile phase with a linear gradient was used: (0 min: 0% B, 9 min: 100% B, 11 min: 100% B, 11.2 min: 0% B, 12 min: 0% B). A UV-VIS detector monitored at 203 nm and the sample injection volume was 2 µl. Standard of gypenoside XVII were isolated by Unhwa. Ginsenoside F2 was purchased from LKT Laboratories.

Microscopy. Light microscopy was undertaken using a model BX41, Olympus. A polarizer for transmitted light, model U-POT, Olympus, was used for TE images. TEs were identified by virtue of their bright fluorescence, due to the presence of lignified secondary cell walls.

Vacuole experimentation was undertaken based on modifications of the methods described previously[44,45]. Briefly, CMCs, needle- and embryo-derived DDCs were stained with 0.01% (wt/vol) neutral red (SIGMA-ALDRICH) for 3 min. Then, cells were washed with 0.1 M phosphate buffer (pH 7.2) and were observed with an optical microscope (BX41 Olympus) using the same buffer.

LC-MS. Analysis was undertaken using an HP 1100 Series liquid chromatography/HP 1100 Series mass selective detector (Agilent Technologies). Samples (2 µl) were separated on a PerfectSil Target ODS-3 (4.6 mm × 150 mm × 3 µm) using water (10 mM ammonium acetate): acetonitrile which was isocratic: 50% acetonitrile for 60 min, at 0.4 ml/min flow rate. Mass detection of paclitaxel was by electrospray ionization (ESI) in the positive ion mode. The drying gas was N_2 at 10 l/min, 350 °C, 30 p.s.i. The vaporizer was set to 300 °C, capillary to 4,000 V. Identification of paclitaxel was accomplished by comparison of retention times and mass with authentic standards.

31. Gamborg, O.L., Miller, R.A. & Ojima, K. Nutrient requirements of suspension cultures of soybean root cells. *Exp. Cell Res.* **50**, 151–158 (1968).
32. Zang, X., Mei, X.-G., Zhang, C.-H., Lu, C.T. & Ke, T. Improved paclitaxel accumulation in cell suspension cultures of *Taxus chinensis* by brassinolide. *Biotechnol. Lett.* **23**, 1047–1049 (2001).
33. Yukimune, Y., Tabata, H., Higashi, Y. & Hara, Y. Methyl jasmonate-induced overproduction of paclitaxel and baccatin III in *Taxus* cell suspension cultures. *Nat. Biotechnol.* **14**, 1129–1132 (1996).
34. Fulcher, N. & Sablowski, R. Hypersensitivity to DNA damage in plant stem cell niches. *Proc. Natl. Acad. Sci. USA* **106**, 20984–20988 (2009).
35. Zhu, Y.Y., Machleder, E.M., Chenchik, A., Li, R. & Siebert, P.D. Reverse transcriptase template switching: A SMART™ approach for full-length cDNA library construction. *BioTech.* **30**, 892–897 (2001).
36. Zhulidov, P.A. *et al.* Simple cDNA normalization using Kamchatka crab duplex-specific nuclease. *Nucleic Acids Res.* **32**, e37 (2004).
37. Margulies, M. *et al.* Genome sequencing in open microfabricated high-density picoliter reactors. *Nature* **437**, 376–380 (2005).
38. Brenner, S. *et al.* Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.* **18**, 630–634 (2000).
39. Schmid, R. & Blaxter, M.L. annot8r: GO, EC and KEGG annotation of EST datasets. *BMC Bioinformatics* **9**, 180 (2008).
40. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
41. Gentleman, R.C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).
42. Bullard, J.H., Purdom, E., Hansen, K.D. & Dudoit, S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* **11**, 94 (2010).
43. Nolan, T., Hands, R.E. & Bustin, S.A. Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.* **1**, 1559–1582 (2006).
44. Kataoka, T. *et al.* Vacuolar sulfate transporters are essential determinants controlling internal distribution of sulfate in *Arabidopsis*. *Plant Cell* **16**, 2693–2704 (2004).
45. Lee, Y. *et al.* The *Arabidopsis* phosphatidylinositol 3-kinase is important for pollen development 1. *Plant Physiol.* **147**, 1886–1897 (2008).

doi:10.1038/nbt.1693